

AD-A061 750

COLORADO UNIV BOULDER CENTER FOR RESEARCH ON JUDGME--ETC F/G 5/10
THE COLORADO REPORT ON THE INTEGRATION OF APPROACHES TO JUDGMEN--ETC(U)
OCT 78 K R HAMMOND, G H MCCLELLAND N00014-77-C-0336
CRJP-213 NL

UNCLASSIFIED

1 OF 5
ADA
061750



DDC FILE COPY

ADA061750



CENTER FOR RESEARCH
ON JUDGMENT AND POLICY

LEVEL 4

11

ADA061750

6 The Colorado Report
on the Integration of Approaches
to Judgment and Decision Making

10 Kenneth R./Hammond, Gary H./McClelland/
Jeryl/Mumpower

12 46pp.

14 CRJP-213

DDC FILE COPY

UNIVERSITY OF COLORADO
INSTITUTE OF BEHAVIORAL SCIENCE

9 Interim rept. Jul 77-Mar 78

DDC
RECEIVED
DEC 5 1978
RECEIVED

Center for Research on Judgment and Policy

Report No. 213

11 Oct 1978

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

15 This research was supported by the Engineering Psychology Programs,
Office of Naval Research, Contract N00014-77-C-0336 Work Unit Number,
NR 197-038. Center for Research on Judgment and Policy, Institute of
Behavioral Science, University of Colorado. Reproduction in whole or
in part is permitted for any purpose of the United States Government.
Approved for public release; distribution unlimited.

78 12 04 242

410 647

Am

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-----------------------|--|
| 1. REPORT NUMBER CRJP 213✓ | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) THE COLORADO REPORT ON THE INTEGRATION OF APPROACHES TO JUDGMENT AND DECISION MAKING | | 5. TYPE OF REPORT & PERIOD COVERED Interim |
| 7. AUTHOR(s) Kenneth R. Hammond, Gary H. McClelland, and Jeryl Mumpower | | 6. PERFORMING ORG. REPORT NUMBER |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Research on Judgment and Policy Institute of Behavioral Science University of Colorado, Boulder, CO 80309 | | 8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0336✓ |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 N. Quincy Street Arlington, VA 22217 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 12. REPORT DATE October, 1978 |
| | | 13. NUMBER OF PAGES 455 |
| | | 15. SECURITY CLASS. (of this report) Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; Distribution unlimited | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Judgment, Decision Making, Psychology, Tradeoffs, Subjective weights | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The present report attempts to take a first step toward the integration of six major approaches to judgment and decision making. These six approaches (and their major spokesmen) can be identified as follows: (a) <u>Decision Theory</u> (Ralph Keeney and Howard Raiffa), (b) <u>Behavioral Decision Theory</u> (Ward Edwards), (c) <u>Psychological Decision Theory</u> (Daniel Kahneman and Amos Tversky), (d) <u>Social Judgment Theory</u> (Kenneth Hammond), (e) <u>Information Integration Theory</u> (Norman Anderson), and (f) <u>Attribution Theory</u> | | |

(E. E. Jones and Harold Kelley). The report attempts to provide a broad, theoretically neutral, systematic framework intended to permit and encourage the discovery of a "kinship system" within which the relationship of each approach to every other approach can be determined.

The report is divided into three major sections, theory, method, and procedure within which each of the six approaches are examined. The theory section deals with the six approaches in terms of their origins, scope, intended function, principal concepts, loci of concepts, and intended use. The method section examines each approach's basic methodology with respect to idiographic versus nomothetic analysis, sampling domains, object decomposition, judgment decomposition, aids for decision makers, and methodological claims. The procedure section focuses on description and comparison of the operational definition that each approach provides for the important concepts identified in the theory section.

In its efforts toward integration the report attempts to achieve the following four goals:

1. denotation of similarities and differences among approaches;
2. specification of gaps and redundancies in the coverage of judgment and decision-making issues;
3. descriptions of antinomies that point to the need for empirical test (crucial experiments) and/or logical or conceptual reconciliation; and,
4. identification of logical and conceptual reconciliation whenever possible, as well as suggestions for empirical reconciliation.

| | |
|---------------------------------|---|
| ACCESSION for | |
| NTIS | White Section <input checked="" type="checkbox"/> |
| DOB | Blue Section <input type="checkbox"/> |
| UNANNOUNCED | <input type="checkbox"/> |
| JUSTIFICATION | |
| BY | |
| DISTRIBUTION/AVAILABILITY CODES | |
| Dist. | AVAIL. AND IN SPECIAL |
| A | |

Table of Contents

| | |
|---|-----------|
| Preface I | viii-ix |
| Preface II | x-xii |
| I. Introduction | I/1-21 |
| Efforts toward Integration | 7 |
| The Present Approach to Integration | 9 |
| Developing a Descriptive Framework | 11 |
| Six Approaches | 13 |
| Decision Theory | 13 |
| Behavioral Decision Theory | 14 |
| Psychological Decision Theory | 15 |
| Transition | 15 |
| Social Judgment Theory | 16 |
| Information Integration Theory | 16 |
| Attribution Theory | 17 |
| A Common Theme | 18 |
| A Descriptive Framework | 19 |
| THEORY | |
| II. A. Comparative Descriptive Section (Theory) | II/A/1-6 |
| Introduction | 1 |
| B. Origins | II/B/1-14 |
| Introduction | 1 |
| Decision Theory | 2 |
| Behavioral Decision Theory | 3 |
| Psychological Decision Theory | 4 |
| Transition | 6 |
| Social Judgment Theory | 7 |
| Information Integration Theory | 8 |
| Attribution Theory | 8 |
| Integration | 9 |
| C. Scope | II/C/1-16 |
| Introduction | 1 |
| Decision Theory | 2 |
| Behavioral Decision Theory | 5 |
| Psychological Decision Theory | 8 |
| Transition | 8 |
| Social Judgment Theory | 9 |
| Information Integration Theory | 9 |
| Attribution Theory | 10 |
| Integration | 10 |

| | |
|--|------------|
| D. Intended Function | II/D/1-37 |
| Introduction | 1 |
| Decision Theory | 1 |
| Behavioral Decision Theory | 4 |
| Psychological Decision Theory | 6 |
| Transition | 9 |
| Social Judgment Theory | 11 |
| Information Integration Theory | 12 |
| Attribution Theory | 16 |
| Integration | 17 |
| Integration of Group I and Group II Approaches | 21 |
| E. Principal Concepts | II/E/11-42 |
| Introduction | 1 |
| Decision Theory | 2 |
| Behavioral Decision Theory | 4 |
| Psychological Decision Theory | 6 |
| Transition | 10 |
| Social Judgment Theory | 11 |
| Information Integration Theory | 16 |
| Attribution Theory | 18 |
| Integration | 22 |
| Integration of Group I and Group II Approaches | 32 |
| F. Loci of Concepts | II/F/1-16 |
| Introduction | 1 |
| Decision Theory | 3 |
| Behavioral Decision Theory | 5 |
| Psychological Decision Theory | 10 |
| Transition | 12 |
| Social Judgment Theory | 12 |
| Information Integration Theory | 13 |
| Attribution Theory | 14 |
| Integration | 15 |
| G. Intended Uses of Research Results | II/G/1-16 |
| Introduction | 1 |
| Decision Theory | 2 |
| Behavioral Decision Theory | 3 |
| Psychological Decision Theory | 4 |
| Transition | 6 |
| Social Judgment Theory | 6 |
| Information Integration Theory | 8 |
| Attribution Theory | 8 |
| Integration | 9 |
| Addendum | 11 |

| METHOD | | |
|---------|--------------------------------------|------------|
| III. A. | Introduction | III/A/1-2 |
| B. | Idiographic vs. Nomothetic Analysis | III/B/1-14 |
| | Introduction | 1 |
| | Decision Theory | 3 |
| | Behavioral Decision Theory | 4 |
| | Psychological Decision Theory | 5 |
| | Social Judgment Theory | 7 |
| | Information Integration Theory | 8 |
| | Attribution Theory | 9 |
| | Integration | 10 |
| C. | Sampling Domains | III/C/1-15 |
| | Introduction | 1 |
| | Decision Theory | 2 |
| | Behavioral Decision Theory | 4 |
| | Psychological Decision Theory | 5 |
| | Social Judgment Theory | 7 |
| | Information Integration Theory | 8 |
| | Attribution Theory | 9 |
| | Integration | 10 |
| D. | Stimulus-Object Decomposition | III/D/1-9 |
| | Introduction | 1 |
| | Decision Theory | 3 |
| | Behavioral Decision Theory | 4 |
| | Psychological Decision Theory | 4 |
| | Social Judgment Theory | 5 |
| | Information Integration Theory | 5 |
| | Attribution Theory | 6 |
| | Integration | 7 |
| E. | Judgment Decomposition | III/E/1-22 |
| | Introduction | 1 |
| | Decision Theory | 1 |
| | Behavioral Decision Theory | 2 |
| | Psychological Decision Theory | 3 |
| | Social Judgment Theory | 3 |
| | Information Integration Theory | 4 |
| | Attribution Theory | 5 |
| | Integration | 5 |
| F. | Methods of Partitioning the Decision | III/F/1-17 |
| | Introduction | 1 |
| | Decision Theory | 5 |
| | Behavioral Decision Theory (SMART) | 8 |
| | Social Judgment Theory | 9 |
| | Integration | 10 |

PROCEDURE

| | |
|--|----------|
| IV. Introduction | IV/1 |
| A. Task Construction/Cue Selection | IV/5-24 |
| Specification by analyst versus judge | 5 |
| Fact/knowledge distinction | 11 |
| Type of response | 11 |
| Judgments of preference or utility | 13 |
| Judgments of inference | 16 |
| Probability estimation | 17 |
| Design | 18 |
| Level of abstraction | 22 |
| Cue metric | 23 |
| Overview | 24 |
| B. Task Characteristics | IV/25-43 |
| Objective Task Characteristics | 26 |
| Uncertainty | 26 |
| Organizing principles | 30 |
| Task elements | 32 |
| Cue-criterion relations | 33 |
| Task element interrelations | 34 |
| Number of task elements | 36 |
| Other task characteristics | 37 |
| Overview | 38 |
| Subjective Task Characteristics | 38 |
| Task element interrelations | 39 |
| Overview | 43 |
| C. Subjective Use of Subjective Data | IV/44-48 |
| Probability estimation | 44 |
| Overview | 48 |
| D./E.F. Subjective Use of Objective Data/ Cue Utilization/Central Processes | IV/48-81 |
| Utility/value/function form | 50 |
| Decision Theory | 51 |
| Behavioral Decision Theory | 54 |
| Psychological Decision Theory | 55 |
| Social Judgment Theory | 56 |
| Information Integration Theory | 57 |
| Attribution Theory | 59 |
| Overview | 59 |
| Weight | 62 |
| Decision Theory | 62 |
| Behavioral Decision Theory | 66 |
| Psychological Decision Theory | 66 |
| Social Judgment Theory | 67 |
| Information Integration Theory | 68 |
| Attribution Theory | 69 |
| Overview | 69 |

| | |
|---------------------------|----------|
| Organizing principles | 71 |
| Bayes theorem | 71 |
| Additive models | 72 |
| Multiplicative models | 75 |
| Multilinear models | 76 |
| Schemata | 76 |
| Overview | 78 |
| Psychological factors | 79 |
| Overview | 81 |
| G. Task-Judge Interaction | IV/81-88 |
| Optimality | 83 |
| Feedback | 86 |
| Feedforward | 87 |
| Overview | 87 |
| H. Conclusion | IV/88 |
| Epilogue | V/1-3 |
| Reference Notes | VI/1 |
| References | VII/1-18 |
| Appendix A | |
| Conference Critiques | |
| Norman Anderson | |
| Ward Edwards | |
| Edward E. Jones | |
| Ralph L. Keeney | |
| Amos Tversky | |
| Appendix B | |
| Distribution List | |

Preface I

The present report was prepared during the period July, 1977 to March, 1978. We undertook its preparation in hopes that it would ultimately contribute to the integration of the numerous approaches to judgment and decision making that have evolved over the past few decades. The more immediate purpose of the report was to provide a common point of departure for a conference that we hoped would mark the start of such integration.

The 12th Annual Conference on Human Judgment was held April 28-30, 1978, in Boulder, Colorado. Twenty-five prominent researchers and theoreticians in the field of judgment and decision making were invited. All but one accepted the invitation, and only one other was unable to attend, a remarkable indication of the interest within the field in the need for integration. A draft of the present report was distributed to all participants prior to the conference; the same report appears here with minor revisions.

The fact that the original draft of this report was prepared as a "working paper" for a small group of individuals who would be called upon to respond and discuss it accounts for several points present readers may find odd or curious. In several places, we adopt a personal tone, entreating the reader to criticize, improve, or otherwise respond to particular sections. We have chosen not to delete these requests for help and assistance simply because

our audience may have expanded. The offers still hold. Also, in numerous places, we inserted the word [MORE] in brackets at the end of a section in order to indicate that we felt there was much left to do at that point. Again, we chose not to delete these indications of incompleteness as our audience increases in size, since we have been unable in the interim to rectify these incompletenesses. The intended nature and function of the report is discussed further in the original preface addressed to conference participants, which is included following this one. The outcome of the conference is discussed in the epilogue to this report.

Work on the integration effort continues. We hope that an improved version of this report will not be long in coming. We wish to take the present opportunity, however, to thank all those whose help has made possible whatever progress we have achieved to date. They include Martin Tolcott (Director, Engineering Psychology Programs at the Office of Naval Research), a number of the conference participants, who provided inspiration and assistance both before and after the conference, and our colleagues and staff at the Center for Research on Judgment and Policy, Institute of Behavioral Science, University of Colorado at Boulder. We would particularly like to acknowledge our gratitude to Lois Stroh, Mary Luhning, Barbara Marvin, and Judy Fukuhara without whose help both the preparation of the report and the conduct of the conference would have far exceeded our capabilities.

Preface II

As will become evident, the present Report is to be considered a "working paper" rather than a completed draft of an article, chapter or monograph. By "working paper" we mean to imply that it is incomplete and inaccurate, and that we enthusiastically invite the reader to work on it--that is, to do whatever s/he can to correct inaccuracies and to take whatever steps s/he can toward completing it. Much of what we do at our conference will depend upon which points you decide deserve priority in discussion, and deserve to have first treatment in the first revision and expansion of what we hope will be "Working Paper No. 2."

In order to underscore the "working" character of this paper, we have noted the sections that are very important and demanding of more work by placing the word [MORE] in brackets at the end of a section. We tried to be judicious about this so that every section would not end in this way, but a lot of them do. We simply had to have some way of conveying to you the fact that we realize that more needs to be done (precisely at that point) and that we didn't have time to do it. Sometimes the reason we didn't have time is that it might have required a week or two or three of reading and thinking before we could have begun to write anything on that matter. Moreover, we often realized that some of our readers had that knowledge (and imagination) at their cognitive finger-tips and could speed up the process by contributing directly

to the completion (or redirection) of that section. Indeed, there might even be two or three different endings provided for a [MORE]. So much the better for now.

Besides our use of [MORE], several other stylistic points should be noted. While all of us have contributed to the ideas and development of each section, the actual writing of a specific section was primarily the responsibility of one person (Section II: Theory--Ken Hammond; Section III: Methods--Gary McClelland; and Section IV: Procedures--Jeryl Mumpower). As a consequence, there are some inevitable style differences among sections even though the basic organization is the same throughout. To highlight that organization and to help the reader skip around in the text, we identify at the top of each page the section (Roman numeral), the subsection (capital letter), and the page number within that subsection (e.g., II/C/26).

All of which brings up the question of the ultimate form of our work. This topic should be discussed at our conference, but certain possibilities are obvious enough to be mentioned now. For example, we can

1. abandon the effort (for the foreseeable future, or for some definite period of time, e.g., x years) for a variety of reasons

2. continue the effort in much the same way. That is, encourage the three present workers (and/or others) to continue to revise and produce Working Paper 2 for revisions at a future conference

- a. a year later
- b. six months later
- c. other

3. produce a Handbook of Judgment and Decision Research (or similar) with a number of contributors, edited by who knows, with an as yet undetermined format, of which the present (and future) working papers might be a part.

- 4. other

Your suggestions will be most welcome.

We hope that you will find that we made a good start.

The Colorado Report on Integration of Approaches
to Judgment and Decision Making

I. Introduction

Contemporary students of human judgment and decision making include operations researchers, management scientists, statisticians, mathematical psychologists and economists as well as others; a wide variety of disciplines have converged on a single topic. What brings these disciplines together is a new vigor in the belief that this critical cognitive activity--judging, deciding, choosing--can be brought under scientific, empirical scrutiny and that as a result of such scrutiny, it cannot only be understood but improved. But the scientific approach to the study of human judgment and decision making is not universally held, or even widely held, to be a method appropriate for this problem; by no means is everyone convinced that science will enable us to understand our cognitive processes. Moreover, the sharp division of opinion as to whether the scientific method is applicable to this problem is age-old, as Raphael shows us in his famous Renaissance painting, in which he arranges the peripatetic philosophers of Athens from right to left in terms of whether they took a quantitative, scientific approach, or an intuitive approach, to understanding nature, including human perception, judgment and thinking.

Even by mid 19th century, these two approaches were still equal contenders for recognition as the proper epistemological method.

Note, for example, how Isaiah Berlin (1977) describes the state of these "rival types of knowledge," in the mid 19th century and how familiar the terms of the dispute will seem today:

The quarrel between these rival types of knowledge--that which results from methodical inquiry, and the more impalpable kind that consists in the 'sense of reality', in 'wisdom'--is very old. And the claims of both have generally been recognised to have some validity: the bitterest clashes have been concerned with the precise line which marks the frontier between their territories. Those who made large claims for non-scientific knowledge have been accused by their adversaries of irrationalism and obscurantism, of the deliberate rejection, in favour of the emotions or blind prejudice, of reliable public standards of ascertainable truth; and have, in their turn, charged their opponents, the ambitious champions of science, with making absurd claims, promising the impossible, issuing false prospectuses, of undertaking to explain history or the arts or the states of the individual soul (and to change them too) when quite plainly they do not begin to understand what they are; when the results of their labours, even when they are not nugatory, tend to take unpredicted, often catastrophic directions--and all this because they will not, being vain and headstrong, admit that too many factors in too many situations are always unknown, and not discoverable by the methods of natural science. Better, surely, not to pretend to calculate the incalculable, not to pretend that there is an Archimedean point outside the world whence everything is measurable and alterable; better to use in each context the methods that seem to fit it best, that give the (pragmatically) best results; to resist the temptations of Procrustes; above all to distinguish what is isolable, classifiable and capable of objective study and sometimes of precise measurement and manipulation, from the most permanent, ubiquitous, inescapable, intimately present features of our world, which, if anything, are over-familiar, so that their 'inexorable' pressure, being too much with use, is scarcely felt, hardly noticed, and cannot conceivably be observed in perspective, be an object of study. This is the distinction that permeates the thought of Pascal and Blake, Rousseau and Schelling, Goethe and Coleridge, Chateaubriand and Carlyle; of all those who speak of the reasons of the heart, or of men's moral or spiritual nature, of sublimity and depth, of the 'profounder' insight of poets and prophets, of special kinds of understanding, of inwardly comprehending, or being at one with, the world. (pp. 78-79)

Any judgment or decision analyst who reads that passage will recognize that Berlin has caught the essence of the antinomy between

the scientific approach to this topic and the "rival type of knowledge," known as common sense or wisdom, so often put forth as superior by the "man on the street." The persistence of the sharp division that Berlin draws between these claims to the correct path to truth that was characteristic of the mid 19th century can clearly be seen in a parallel description of the present rivalry between "common sense" and "refined knowledge" drawn by Stephen Pepper in 1948:

Tension between common sense and refined knowledge.--This is a strange set of traits for an important mass of cognitive material [common sense]--to be not definitely cognizable, to be not cognitively responsible and so irritable, and yet to be cognitively secure. The first two traits in the order just indicated are negative in the eyes of knowledge; only the last is positive. The first two traits are, in fact, so displeasing to experts of cognition that the material of common sense has very frequently been ignored as a respectable factor in cognition. And so, on this side, we find common sense, the opinions of the man in the street, disparaged and ridiculed in comparison with the definite and responsible knowledge of science and philosophy. Yet on the other side, the security of common sense does not wholly escape the attention of men, nor can men wholly ignore an insecurity in the abstract concepts, the hairsplitting definitions, the speculative hypotheses of expert critical knowledge. So, on this score, common sense becomes an object of praise for its simple homespun wisdom and plain practical sense.

This tension between common sense and expert knowledge, between cognitive security without responsibility and cognitive responsibility without full security, is the interior dynamics of the knowledge situation. The indefiniteness of much detail in common sense, its contradictions, its lack of established grounds, drive thought to seek definiteness, consistency, and reasons. Thought finds these in the criticized and refined knowledge of mathematics, science, and philosophy, only to discover that these tend to think out into arbitrary definitions, pointer readings, and tentative hypotheses. Astounded at the thinness and hollowness of these culminating achievements of conscientiously responsible cognition, thought seeks matter for its definitions, significance for its pointer readings, and support for its wobbling hypotheses. Responsible cognition finds itself insecure as a result of the very earnestness of its virtues. But where shall it turn? It does, in fact, turn back to common sense, that indefinite and irresponsible source which it so lately scorned. But it does so, generally, with a bad grace. After filling its

empty definitions and pointer readings and hypotheses with meanings out of the rich confusion of common sense, it generally turns its head away, shuts its eyes to what it has been doing, and affirms dogmatically the self-evidence and certainty of the common-sense significance it has drawn into its concepts. Then it pretends to be securely based on self-evident principles or indubitable facts. If our recent criticism of dogmatism is correct, however, this security in self-evidence and indubitability has proved questionable. And critical knowledge hangs over a vacuum unless it acknowledges openly the actual, though strange, source of its significance and security in the uncriticized material of common sense. Thus the circle is completed. Common sense continually demands the responsible criticism of refined knowledge, and refined knowledge sooner or later requires the security of common-sense support.

Why cannot the two merge? No doubt, that is the inherent aim of cognition. For what the question amounts to is, Why is there any ignorance? It is clear that the answer to such a question can only be given with any specificity in terms of refined knowledge. . . . It seems fairly obvious that as long as refined knowledge is not complete, so long at least will there be a discrepancy between the material of common sense and that of critical cognition. For, considering the situation at its worst, even the extremest efforts of dictatorial propaganda cannot stop those insistent questionings that well up in the most innocent as also in the most sophisticated minds.

Whence do these questionings well up, which are the signs of the obstinate security of common sense? For though man reason himself into a machine, into a solipsism of the present moment, into Nirvana, or into Nothing, life still breaks out in hunger and craving, and nature affirms itself in the strong pressure of the ground and the heat of the sun. There is no doubt of these common-sense insistences, but if we seek the reasons for them we can find them only in refined critical knowledge. . . . But until ignorance completely disappears we cannot expect a specific and fully adequate answer.

Such, then, is the basic polarity of cognition, which we may expect to continue as long as we fall short of omniscience. On the one side, irresponsible but secure common sense; on the other, responsible but insecure critical cognition. (pp. 44-47)

These passages from students of the history of thought are quoted at length because, despite their difference in style, they will convey to the reader a perspective not ordinarily encountered; they illustrate the fact that the study of human judgment and decision making is not a new idea, that the rivalry between different views of what knowledge

is has a long history, and that this rivalry may be expected to persist "as long as we fall short of omniscience," as Pepper puts it. Moreover, these passages should serve to indicate to the reader that the pursuit of the study of human judgment and decision making will lead one to encounter one of the fundamental problems of mankind: How do we know our world and choose what to do? And by what method shall we find the answer?

The longer view aside, however, it will be noted that Pepper's remarks were made in 1948; the reader will now want to know: is the situation no different in 1978? Has the scientific approach not filled out to any significant degree the "thinness and hollowness of conscientiously responsible cognition"? Not yet "found matter for its definitions"? Nor "significance for its pointer readings"? Nor support for its wobbling hypotheses"? In short, does critical cognition not yet offer a sense of security as well as responsibility?

The modern, quantitatively oriented descendants of those on the right of Raphael's painting who study human judgment and decision making will answer in the affirmative: they will say that responsible cognition, refined knowledge, has indeed become more secure, if not fully so. The greatest source of increased security comes from the increased reliance on the use of the logic of mathematics and the resort to empirical test of hypotheses; greater rigor in thought, method, and empirical research has led to more cognitive security. As a result, even though they are aware of how "thin" and "hollow" the results of their efforts can be, and even though the security of common sense is not altogether despised, common sense is under sharper scrutiny (and sharper attack) than ever before by the advocates of responsible

cognition. It has been the (slow and uneven) cumulative growth of hard-won empirical knowledge and skill that has increased the security that responsible cognition has sought, and found, for its "wobbling hypotheses." For while the mid 19th-century philosophers were continuing to pursue the nuances of meaning in the verbal struggle between "rival types of knowledge," mid 19th-century scientists, little known then and now, were successfully making quantitative analyses of judgments for the first time, and were thus preparing the way for the mid 20th-century judgment and decision analysts. The fact that judgments can be quantified, can be refined in a responsible way, and can be analyzed and understood in quantitative terms, can no longer be doubted.

And that brings us to the reasons for this monograph. Awareness of the possibility of mathematically-oriented, empirically-based methods for analyzing and perhaps improving common sense has led to a proliferation of quantitative concepts and techniques, and these are now so numerous and diverse they constitute a threat to the cumulative nature of the scientific study of judgment and decision making; the proliferation of effort is so large and so scattered over the academic and nonacademic terrain that its very profuseness is a danger to its future. It will, therefore, be our purpose to attempt to organize and integrate contemporary approaches to this topic in order to assist in the continuance of the cumulative growth of what may, because of its central function, become a separate, identifiable scientific discipline.

Efforts Toward Integration

The present need for integration is reflected in the fact that at least three sponsored conferences and at least nine articles (not counting review articles) have recently addressed themselves to the problem of the proper study of judgment and decision making.

The following two paragraphs from the introduction to one recent conference indicate the breadth of the academic interests of the participants:

The search for theories, concepts, and techniques applicable to decision-making processes under multiple criteria has steadily been intensified during the past few years. We feel that the time has come to review the results, to compare them, and to outline possible future directions. As far as we know, the seminar at which the papers in this volume were read was the first meeting of this nature in the United States. To assure the highest quality of this meeting, we have sought participation of leading researchers and practitioners in this field. Since the problem is highly general, the participants represent many areas of applied decision making. Most are from business and economics, but a number of engineers, psychologists, sociologists, behavioral scientists, mathematicians, statisticians, and political scientists also contributed (*italics ours*).

The important part of the discussions explored the relation between formalized decision-making techniques (utilizing computer and mathematical analysis) and of human judgment, based on intuition, experience, and "professional insight." Rejuvenation of the role of human judgment seems to be one of the main aspects of the literature on multiple criteria decision making but many participants seem to be skeptical about man's ability to choose among multiattributed alternatives. (Cochrane & Zeleny, 1973, p. xiv)

There are few signs, however, that paper-reading conferences lead to integration. And although it is too soon to appraise the result of the most recent integrative efforts (Fischhoff's "Attribution Theory and Judgment Under Uncertainty," 1976, and Shanteau & Phelps' "Judgment and Swine: Approaches and Issues in Applied Judgment Analysis," 1977), the results of similar articles are discouraging;

progress seems not to have occurred. For example, Anderson's effort in 1974 to integrate the approaches of Information Integration Theory and Attribution Theory can only be described as heroic, yet in a recent compendium on Attribution Theory (Harvey, Ickes, & Kidd, 1976) that includes 26 authors, Anderson's article is mentioned only once, and that reference ignores his attempt at integration. Similarly, Kukla's effort (1972) to "combine concepts of decision theory with those of attribution theory" (p. 454) is not mentioned in that compendium by a single author (except Fischhoff).

Nor is there any evidence that the other efforts at integration have achieved recognition, or have changed anyone's approach to the study of judgment and decision making. Typically, related work carried out by a member of a different discipline is given an academic "nod." Keeney and Raiffa's footnote (p. 8) regarding the contributions of psychologists provides a good example, thus: "Clearly, there is much overlap of interest between the prescriptive and descriptive approaches. Over the past 25 years, the contributions of many people concerned with the descriptive analysis of decision making have had a significant impact on prescriptive decision analysis." Keeney and Raiffa then refer to four "excellent reviews." But they fail to indicate what these "significant impacts" are, and reading Keeney and Raiffa did not enable us to discern them; even Edwards' work on Bayesian decision making is given only passing mention. The only hopeful note we find is in (alas!) another footnote (p. 212) in which Keeney and Raiffa acknowledge that "Tversky and other experimental psychologists working in descriptive decision theory indicate that assessors inadvertently bias responses by the form of

the questions they pose," an acknowledgment that can hardly be said to capture the full spirit of the work done so far.

Not to be one-sided, we must also note that the academic "nod" relationship is symmetrical. While psychologists studying judgment and decision making sometimes cite the precursors of modern decision theory (e.g., von Neuman & Morgenstern and Savage), one must look hard to find any serious consideration by psychologists of recent work in decision theory. In short, it is not that bridges need to be built between different approaches, rather, a kinship system must be found within which each approach can determine its relationship to every other approach.

The Present Approach to Integration

We shall, therefore, try to provide a broad, theoretically neutral, systematic framework that will permit, indeed, encourage the discovery of that "kinship system" rather than to compare pairs of approaches in detail. Our framework will be broad because the contributions to the study of judgment and decision making come from a wide range of theoretical approaches. Only a broad framework that encompasses these apparently disparate approaches can help us to determine whether there is a common core of concepts, methods and procedures within them, and what the boundaries of that common core might be.

Our framework will attempt to achieve theoretical neutrality by comparing each approach on a set of neutral dimensions that will be indicated below. Theoretically neutral description should make systematic comparison possible and, we hope, acceptable as well. Systematic treatment of several approaches over a series of dimensions

has the virtue of demanding completeness from the author and providing it for the reader. But just as we have learned that it requires a patient, even pedestrian, diligence on the part of the authors, the reader will learn that s/he will have to be patient, too. Good organization of what is generally known about various approaches, accuracy and completeness of description of these approaches--these are our aims, not the overwhelming insight that produces yet one more new approach, or that declares that only one of the present approaches is on the narrow road to truth (which, we admit, remains indistinct to us).

Our work is programmatic as well as systematic. We have no illusions that our overview will be the end of this project; it is clearly only the beginning. We shall take full responsibility for providing a good organizational framework that can be used, if only to be used to develop a better one. Failure here will be wholly our failure. We do not take full responsibility for accuracy, however; we shall simply do the best we can. The knowledge that our descriptions and interpretations of the work of others will be reviewed and criticized by them, however, gives us full confidence that few inaccuracies will fail to be detected and corrected with precision, perhaps even enthusiasm.

Completeness, of course, is far beyond our capacity: it is a goal that can only be approximated more closely over time. That is why our work is programmatic. Our resources permitted us to do no more than rely on a few major sources, although these are generally original ones. Here again we rely on the personal interests of the

proponents of various approaches to provide the materials as well as the guidance to make certain that first things are dealt with first. But the reverse is true as well; conceptual boundaries must be drawn; not everything that a proponent of a given approach has done or written is relevant to this project. Therefore, while errors of fact and interpretation can be corrected rather quickly, incompleteness can be remedied only through contributions to a programmatic effort by all concerned.

It is precisely this programmatic aim that most significantly marks off the present method for achieving integration from previous ones. We shall present a framework for integration that invites various researchers in the field of judgment and decision making to participate in the persistent reconstruction of that framework and the continuous alteration of its content. Thus we invite researchers to reconstruct that framework in a form that brings integration closer.

Our approach, then, will be to provide a broad, systematic descriptive framework for the reader to improve, in addition to a report to contemplate. We hope that our readers will find our report to be sufficiently well organized and accurate to be worthy of their efforts to complete and improve it, and thus to participate in the initial step toward the development of a cumulative scientific discipline of judgment and decision making.

Developing a Descriptive Framework

Theoretical approaches cannot be integrated without describing and comparing them, and, therefore some means must be found for making comparative description possible, as well as palatable.

Because we chose to develop a broad descriptive framework ours encompasses approaches from decision theory (as it is developed and applied in the field of management science) to attribution theory (as it is developed and applied in psychology). Six approaches were included within that range:

Decision Theory (DT)

Behavioral Decision Theory (BDT)

Psychological Decision Theory (PDT)

Social Judgment Theory (SJT)

Information Integration Theory (IIT)

Attribution Theory (AT)

It will be immediately apparent to students of judgment and decision making that we took some liberties with the names we assigned to the second and third approaches. We have deliberately and (somewhat) arbitrarily assigned the name of Behavioral Decision Theory to the work of Edwards and his colleagues, and the name Psychological Decision Theory to the work of Kahneman and Tversky and their colleagues (among whom we include Fischhoff, Lichtenstein, and Slovic). Our defense is simply that we believed the distinction to be necessary, a defense the reader will have an opportunity to evaluate below.

Although others may well have chosen to consider other, or more, approaches, we believe that an effort to integrate these six approaches should satisfy the requirements of an initial effort. True, we have cast a net with a very broad mesh. But the use of a finer mesh, the inclusion of more approaches, would have meant a long postponement of this report and criticism of it. As it is, the reader may well wish that we had cast even a broader net, for then we could have more nearly completed our work.

Six Approaches

The six approaches to be described in this monograph include three whose origins lie primarily in economics (Decision Theory, Behavioral Decision Theory, Psychological Decision Theory) and three whose origins lie primarily in psychology (Social Judgment Theory, Information Integration Theory, and Attribution Theory). Each is discussed very briefly below.

Decision Theory (DT)

Of a number of possible approaches within Decision Theory we chose to describe the work of Ralph Keeney and Howard Raiffa not only because of their steady series of contributions to this problem, but because of the recent organization of their work into a well-recognized book (Keeney & Raiffa, 1976) that provides a highly accessible reference work. When we refer to DT in the remainder of the text, therefore, we shall be referring to that version of DT that is contained in their book, a version often referred to as "Multiattribute Utility Theory."

Main topic. The main topic for DT is the problem of choosing among alternatives with multiple attributes. According to DT, the principal parameters of the process of choosing this alternative over that one are (a) the probability of the occurrence of the alternative and (b) its utility to the decider. This formulation of the decision process lends itself to formal, mathematical analysis, thus enabling the decision theorist to address a fundamental question: does the decision process of this (or any other) decision maker conform with the formal axioms of choice, as set forth by the logic of the mathematics of choice? Having found evidence that the cognitive activity of decision

making often does not meet this criterion, decision theorists have developed means to assist decision makers to achieve rationality by bringing their cognitive activity in conformity with the demands of logic. DT is, then, essentially a means for prescribing what the decision process should be, if it is to be rational. Decision analysis, the art of assisting decision makers, as well as research on decision theory, are now thriving activities.

Behavioral Decision Theory (BDT)

The discovery that human decision making falls short of meeting the demands of rationality was hardly a new idea, and it was not greeted, inside or outside the academic world, with astonishment, or even doubt. But a new vigor was given to the study of the cognitive aspects of decision making when a psychologist--Ward Edwards--observed the progress that had been made by economists, noted the parallel interests of psychologists in the topic of decision making, showed how the work of psychologists and economists could supplement one another, and, most important, showed how supplementation could produce a systematic empirical program of research.

Main topic. In particular, Edwards focused on the question of how to describe the less-than-optimal behavior of the decision maker, and offered a theory of what the behavior of the decision maker is (as compared with the decision theorists' interest in showing what it is not). Edwards' effort to document the value of Bayesian statistics as a descriptive point of reference led to a number of similar efforts by others, nearly all of which were tested in the psychological laboratory, within the general framework of psychological research

methodology, as well as with due regard for the theory of psychological measurement that had begun over a century earlier. In addition, Behavioral Decision Theorists have also investigated the value of Multiattribute Utility Theory as a method for assisting decision makers, in a manner analogous to that of Keeney and Raiffa.

Psychological Decision Theory (PDT)

Roughly 20 years after Edwards' classical 1954 paper, two psychologists, Daniel Kahneman and Amos Tversky, took the psychological approach further; they developed a theory of decision making that moved beyond description to explanation and prediction of decision behavior.

Main topic. The main interest of these researchers is to find the cognitive sources of the departure from the criteria of rationality. Specifically, they seek the manner in which the processes of memory, perception and specific varieties of experience lead decision makers to develop systematic errors in their estimates of the probabilities and utilities that are the key parameters in decision theory.

Transition

In short, DT is largely if not entirely prescriptive in its efforts, whereas the latter two approaches are descriptive--and explanatory. That is, whereas DT indicates how the decision maker should arrive at rational decisions, the latter approaches describe how and why they depart from DT's rational prescription for cognition that include judgments of probability and utility. The approaches we shall next describe, however, are wholly concerned with the description of the process of judgment, without significant regard for prescription. And that is because they are primarily concerned

with how people know, rather than how they choose, a distinction that is important, not only because it specifies differences in aims, but because it carries implications for the methodology that researchers use.

Social Judgment Theory (SJT)

This approach has its origins in the field of perception, and takes its general theory and method from the work of Egon Brunswik (1903-1955). Kenneth Hammond extended this approach to the study of judgment and decision making, and, in doing so, continued to emphasize the cognitive difficulties created by the probabilistic, interdependent relation among variables in the world (i.e., the causal texture of the natural environment) as it is experienced by the person attempting to understand it. SJT is thus related to the approaches described above because it shares their emphasis on the uncertainty inherent in the physical, biological and social environment.

Main topic. SJT gives considerable attention to the manner in which the formal properties ("causal texture") of the environment create significant difficulties for human beings to learn to make accurate judgments about environmental events (multiple cue probability learning) including the behavior of other people (interpersonal learning). Interpersonal conflict arising from different judgments is also a topic to which SJT gives considerable attention. In general, SJT emphasizes the interaction between environmental and cognitive systems.

Information Integration Theory (IIT)

There is a direct line from the work of the psychophysicists mentioned earlier to the approach taken by Norman Anderson in his

development of IIT. For in their efforts to understand the cognitive process of integrating information, IIT theorists place considerable emphasis on measuring, in a precise metric sense, the social as well as physical judgments of persons.

Main topic. IIT emphasizes and provides for the analysis of the cognitive integration of multiple pieces of information that are measured subjectively, and for which subjective importance is also measured. A person's method of integrating such information is indicated in terms of the "cognitive algebra" employed. The principal aim of IIT is to discover the form of cognitive algebra human beings employ in various cognitive activities required by various tasks. For example, is the information in one set of conditions integrated or combined by one form of cognitive algebra (e.g., averaging), while information in other conditions is integrated by a different algebraic form (e.g., multiplying)?

Attribution Theory (AT)

The psychology of knowing in social as well as physical circumstances was emphasized by Fritz Heider, Brunswik's contemporary. Just as Brunswik emphasized the "causal texture" of the environment, Heider emphasized the cognitive difficulties of attributing causes when they are buried in the structural interdependency of the variables in the environmental context.

Main topic. Causal attribution has been the principal topic of interest to researchers following Heider's general approach to inference. Under what conditions will a person be seen as a causal agent in contrast to an environmental circumstance? When will a person rather than circumstances be blamed (or rewarded) for outcomes

of events? Causal attribution is of course, a central feature of "the psychology of common sense" and, therefore, Attribution Theory addresses directly the "tension between common sense and refined knowledge" of which Pepper wrote. Edward Jones and Harold Kelley, as well as many others, have pursued these questions primarily by means of the experimental approach.

A Common Theme

For those on the left of Raphael's Renaissance painting, the wisdom of common sense is a combination of rationality tempered by experience, a powerful form of cognition not to be destroyed by putatively scientific methods of analysis that decompose it into artificial and "hollow" fragments. For those quantitatively oriented philosophers on the right, however, human cognition is not beyond analysis, nor does analysis necessarily destroy it. Since cognition is a natural activity, it requires scientific understanding, and should be susceptible to the empirical method that is basic to all science. But those who intend to bring the scientific method to bear on the process of human judgment and decision making know that their work will be protested (even by scientists), not merely because of its boldness, but because of its arrogance. For by asserting their premise--that human cognition can and should be studied by scientific methods--they imply that they expect not only to understand one of nature's most mysterious processes, but expect to control it as well. And therein lies the protest of the 20th century against the effort to "calculate the incalculable," as Berlin phrased it. But as the remainder of this monograph indicates, that protest has failed to discourage scientific efforts to understand human judgment and decision processes.

We turn now to the development of a systematic, descriptive framework for comparing and integrating these scientific efforts.

A Descriptive Framework

In our effort to be systematic, we began with fundamentals; we divided our work into three major sections, theory, method, and procedure, and examined the six approaches mentioned above accordingly.

The theory section deals with the six approaches in terms of their origins, scope, intended function, principal concepts, loci of concepts, and their intended use. The method section examines each approach's basic methodology with respect to idiographic versus nomothetic analysis, sampling domains, object decomposition, judgment decomposition, aids for decision makers, and methodological claims. The procedure section focuses on description and comparison of the operational definitions that each approach provides for the important concepts identified in the theory section.

Having indicated the purpose of integration--the development of a cumulative scientific discipline--and the general method we employ in order to initiate this effort, we turn to the question of what integration could possibly mean. We take it that integration requires the following steps:

1. denotation of similarities and differences;
2. denotation of gaps and redundancies in the coverage of judgment and decision-making issues;
3. denotation of antinomies that point to the need for empirical test (crucial experiments) and/or logical or conceptual reconciliation;

4. logical and conceptual reconciliation whenever possible, as well as suggestions for empirical reconciliation, although we hasten to add that integration does not mean "amalgamation," "coalescence," or even "unification." Our purpose is not to eliminate differences, providing that they are not merely verbal, or trivial in other ways.

And what if these steps were accomplished to some detectable degree? What would have been accomplished? First, a strong move would have been made toward the intellectual coherence that is now lacking; second, a clearer understanding of what various approaches add to the growth of scientific knowledge in the area of cognition would have been achieved, and third, a greater appreciation of the ability of various approaches to contribute to the judgment and decision-making capacity of human beings.

Are these important tasks? To ask that question of our readers may be analogous to asking the workers in Newcastle if coal is important. But it may not be superfluous to point out just how far we have already come in our own judgments of the validity and significance of our work. Consider these remarks made by Paul Slovic in the context of an article addressed to laymen: "This work [on judgment and decision making] has led to the sobering conclusion that, in the face of uncertainty, man may be an intellectual cripple (italics ours), whose intuitive judgments and decisions violate many of the principles of optimal behavior" (1976). If there was tentativeness in this "sobering conclusion" conveyed by the word "may," that tentativeness was obliterated in the next sentence: "These intellectual deficiencies underscore the need for decision-aiding techniques" (p. 222). Slovic's views are

undoubtedly shared by other scientifically-oriented students of judgment and decision making. Are these the conclusions that the philosophers on the right of Raphael's painting would have anticipated?

T H E O R Y

A. Comparative Descriptive Section (Theory)

Introduction

This section addresses itself to two problems: (a) describing and comparing the theoretical content of the six approaches to judgment and decision theory, and (b) attempting to integrate those approaches.

We begin our description and comparison of the six approaches by ordering them along a continuum as indicated below:

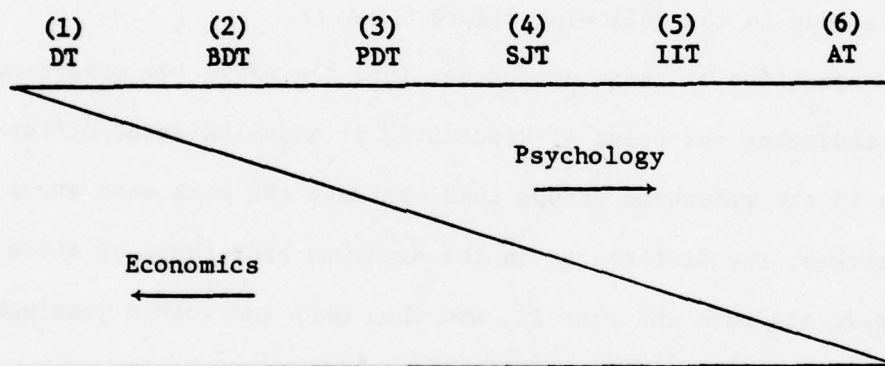


Figure II-A-1

Any such single-dimensional ordering will of course be unsatisfactory. The above ordering will, however, serve the purpose of indicating the origins, or point of departure, of the six approaches. It represents roughly the largely economic reference of the approaches on the left, a reference that gradually diminishes as we move to the right on the continuum, and

concomitantly represents the gradual increase in emphasis on psychology. This change of emphasis is also reflected in the fact that in all of the three approaches on the left (1, 2, and 3), concepts of probability, utility and aggregation are taken to be major ones, whereas the three approaches on the right (4, 5, and 6) treat these concepts as subvarieties of larger concepts of interest. Consequently, we shall categorize the first three approaches as Group I, and the second as Group II. As we shall see, however, integration is as likely to be needed within these groups as between them.

The separation of the six approaches into two major groups can be seen in the following figure (II-A-2).

Separation of these approaches into the above two groups not only indicates our point of departure, it also indicates differences in the reference groups that evaluate the work each group undertakes, the differences in the academic background of those who read the work and cite it, and thus make the work a prominent (or obscure) part of the literature.

Having indicated the manner in which the six approaches were ordered and categorized we now attempt to describe them in terms of the following theoretically neutral dimensions:

- B. origins of approaches (e.g., economics, psychology)
- C. scope of approaches (e.g., single person decisions, interpersonal learning)
- D. intended functions of the theories (e.g., descriptive, explanatory)

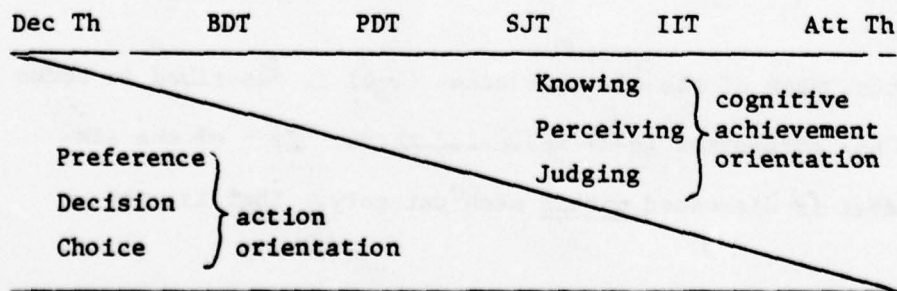
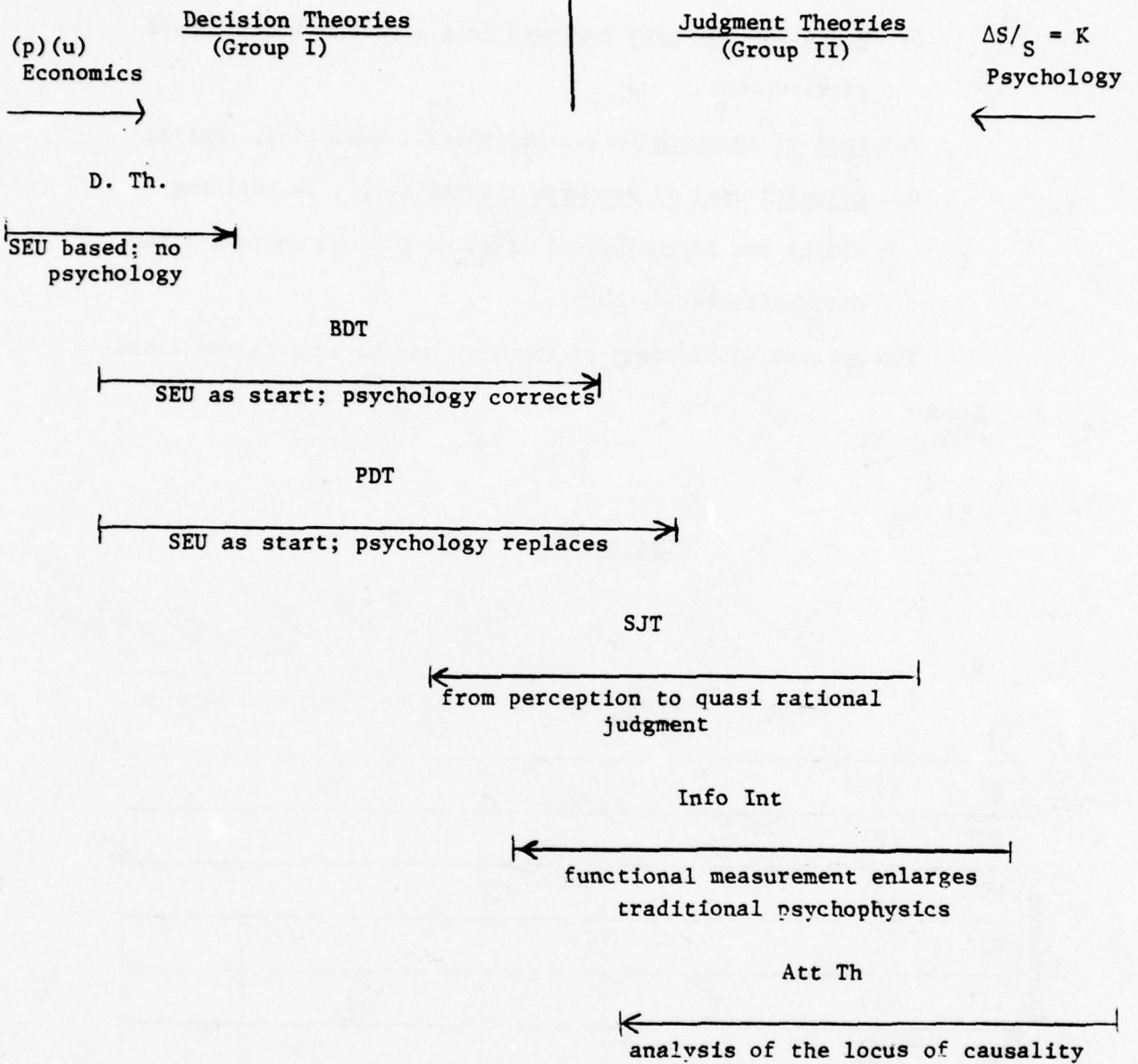


Figure II-A-2

- E. principal concepts employed (e.g., probability, causal attribution)
- F. loci of concepts (e.g., peripheral, mediating, central)
- G. intended uses of research results (e.g., as building blocks for psychological laws, or for use in real-life circumstances, or both).

The general development of the work may be seen in the table below:

| | | Categories | | | | | |
|------------|---------|------------|-----------|--------------------|--------------------|------------------|-------------------------|
| | | Origins | Scope | Intended functions | Principal concepts | Loci of concepts | Intended uses of theory |
| | | Section B | Section C | Section D | Section E | Section F | Section G |
| Approaches | DT (1) | | | | | | |
| | BDT (2) | | | | | | |
| | PDT (3) | | | | | | |
| | SJT (4) | | | | | | |
| | IIT (5) | | | | | | |
| | AT (6) | | | | | | |

Table II-A-1

Thus, each of the six approaches (1-6) is described in terms of the six categories (B-G) indicated above. Each of the six approaches is discussed within each category. That is, the

discussion proceeds down successive columns of the table, thus providing a discussion of DT in terms of its origins, the names of its primary concepts, its scope, etc. BDT is then discussed in terms of the same six categories, followed by a similar discussion of the remaining four approaches. Admittedly far more general than previous efforts at integration, we believe the loss in specificity is compensated for by the gain in systematization. Systematic treatment of the six approaches makes comparison as well as description possible, and thus brings integration within our reach, if not our grasp. Future work can bring more specificity to the comparisons.

Integration is attempted by two means: first, a "Transition" section appears between the discussions of the Group I and Group II approaches. The purpose of the Transition section is to remedy to some extent the inaccuracy we create by the division of the six approaches into two groups, when there are, in fact, many similarities between them (as suggested in Figure 2). That is, the Transition section describes the conceptual watershed, not the conceptual barrier, that lies between these two general approaches and thus indicates the diffuse rather than sharp differences between approaches on either side of the division.

The second means for achieving integration is more direct; a section at the end of each of the six descriptive categories (origins, intended function, etc.) is entitled "Integration." It is this section, of course, that is the most important, because

it is here that we attempt to integrate the material from the six approaches with regard to each of the categories indicated above.

The reader may accept our definition of the task of integration yet still be curious about our conception of a theory, conspicuously absent so far. We have not attempted to present a set of criteria that must be met before an "approach" could achieve the status of a theory because we are not analytical philosophers. Nor do we believe that the approaches to judgment and decision making have progressed to the point where such analytical treatment would prove useful. Rather, we simply described each approach in terms of the categories indicated above. We recognize that clear and concise definitions of "similarities," "differences," and "antinomies" require an analytical treatment of the question of what the form and substance of a theory in the field of judgment and decision behavior should consist of. All in good time.

We turn now to a discussion of each of the six approaches in terms of the six descriptive categories indicated in Table 1.

B. Origins

Introduction

Why should a paper directed toward the integration of different approaches to the study of judgment and decision making begin with comments on the origins of these approaches? Because (a) the theories, methods and procedures used by different approaches, (b) the judgment and decision aids they offer to those who want them, and (c) the criteria by which all of these are evaluated as "right," "wrong," "competent," "useless," "correct," "misguided" stem directly from their origins. And unless the story of the development of standards as well as approaches is unraveled, there is little likelihood of the development of a cumulative scientific discipline.

Obviously, we cannot even begin to write the complete history of the development of these approaches. We should, however, at least point to the different origins of different approaches; otherwise integration would be impossible.

The study of judgment and decision making has two primary sources--economics and psychology. And mathematics hovers above, beyond, or around them, thus providing the logical context for the study of judgment and decision. In economics, people are believed to decide on a rational course of action (to buy, to sell, in general to accept or not accept an alternative) because they have values (utilities) and beliefs about the likely occurrence of events (probabilities), giving rise to what was labelled

the "rational man" theory of consumer choice. That is enough for (at least some) economists, who then explore the consequences of the choices and actions that follow from various sets of utilities and beliefs. Psychologists (and philosophers), on the other hand, have wanted to know what are the sources of beliefs, expectancies and preferences rather than merely to explore the consequences of holding them.

We do not, of course, recount here the history of the growth of the interaction and gradual merging (and dim but certain recognition of the merging) of these general approaches (nor the acknowledgment by modern philosophers of advances in judgment and decision analysis, as for example, in Rawls, 1971 or Hawkins, 1977), but readers will be aided if they keep these different origins in mind. For not only do different concepts emerge from these different origins, but different aims, methods and procedures emerge as well.

Decision Theory

The mathematical terms of DT (probabilities, utilities, and methods of aggregation) were taken from the "behavioral-mathematical" theory called "expected utility theory" constructed by economists (Morgenstern) and mathematicians (von Neumann). Expected utility theory is itself a derivative of an earlier utility theory (the value theory of the 19th-century economists), and can also be traced to the famous Bernoulli expression that indicates that the worth of a decision is determined by the probability of events and their associated utilities.

In short, the major concepts of DT are a legacy from the 19th-century economists, the principal modifications being (a) the acknowledgment in the 20th century of the presence of uncertainty in the mind of the decision maker, and (b) the effort to deal with it directly. The implications of these concepts have been refined by the mathematical logic of the 20th century and made available to decision makers by the work of such decision theorists as Keeney and Raiffa and others. A guide to the development of Multiple Criteria Decision Making (not to be confused with MAUT) can be found in Starr and Zeleny (1977).

Behavioral Decision Theory

Behavioral decision theory (BDT) is the brainchild of Ward Edwards, who coined this term as well as "Subjective Expected Utility" (SEU). The origin of BDT is clearly set forth in his landmark article "The Theory of Decision Making," (reprinted in Edwards' and Tversky's "Decision Making," Penguin, 1967, to which page numbers in this paper will refer). Although Edwards was not the first psychologist to become interested in this topic, the 1954 article gave it a coherence that enabled psychologists to see that an important type of behavior needed their attention, and that its study could be interesting and rewarding. There is little doubt that this article and his later one in the Annual Review of Psychology in 1961 will become classics in the history of research in judgment and decision making. These seminal articles include numerous references

to economic theory, (economic) decision theory, game theory, clinical psychology, psychophysics, probability and statistics, and psychometrics. From this point of departure, Edwards developed an interest in the Bayesian theory of statistics, an interest that led to a large number of studies by Edwards and, very soon, his numerous colleagues, including Peterson, Phillips, Beach and others. Bayesian research then led to an interest in multi-attribute utility theory (MAUT) in the 1960s and thence to Edwards most recent technique for aiding public policy decision makers—SMART (simple multi-attribute rating technique; see Gardiner & Edwards in Kaplan & Schwartz, 1975).

BDT now encompasses more than Edwards work, of course, and as may be seen from the 1977 Annual Review of Behavioral Decision Theory, work in this area has grown considerably. The 319 references cited by Slovic, Fischhoff and Lichtenstein, 1977 probably constitute less than one half of the articles and books written on the general topic within the period considered. In general, research in this area has moved toward increasing the role of explanatory psychological concepts. And that effort has developed such a large momentum that we have grouped these studies under the separate heading of Psychological Decision Theory (PDT).

Psychological Decision Theory

Although Slovic and Lichtenstein (1971) were able to review a few information-processing studies of judgment and decision making, the recent interest in PDT clearly originated in the heuristics studies of Amos Tversky and Daniel Kahneman—regarding the "law

of small numbers" (1971), "representativeness" (1972), and "availability" (1973). These studies marked a sharp change from previous investigations of judgments of uncertainty. Earlier studies (mostly within BDT) had concentrated on descriptions of subjects' judgment and decision behavior in terms of (or as deviations from) basic normative models. Tversky and Kahneman rejected further "baseline" descriptions of judgment and emphasized the search for the psychological (mainly cognitive) mechanisms which people use to evaluate frequencies and likelihoods. The reason for the change in emphasis (from behavioral to psychological DT) was apparently that there had accumulated sufficient empirical generalizations (indicating that deviations of subjective from objective probability were reliable, systematic, and difficult to eliminate) to demand a "systematic theory about the psychology of uncertainty" (Kahneman & Tversky, 1972, p. 430).

Before concluding this section, we should note that the uninitiated (and others) may be excused if they should become confused about the use of terms such as SEU and MAUT. The origin of the term SEU is clear enough; Edwards in 1961 stated that: "Work since [1954] has focused on the model which asserts that people maximize the product of utility and subjective probability; I have named this the subjectively expected utility maximization model (SEU model)" (Edwards & Tversky, 1967, p. 67). But it appears that this appellation has not been accepted by those to whom it was intended to apply. It was not mentioned by Raiffa in 1968, by Fishburn in 1972, nor by Keeney and Raiffa in 1976.

Multi-attribute utility theory is, however, discussed by Raiffa in 1968 and the reader is referred to several sources, mainly in the 1960s, for further reading.

The use by Edwards of both Bayesian concepts and MAUT can be seen in his article in Hammond, 1978: "Judgment and Decision in Public Policy Formation."

Transition

Moving from Group I approaches to Group II approaches shifts the origins of theoretical concepts. As noted above, Group I has its origins in the concepts introduced by economics (primarily in the theory of consumer choice), but Group II has its origins in psychology (largely in the psychology of perception).

The general theme of Group II ignores the motivational component so central to Group I (people try to "maximize something"). Although Group II retains the concept of "intention," this concept has a meaning more nearly akin to "focus" than to "getting." Thus, the study of "knowing" rather than "getting" is the main theme of Group II.

Although the origin of the concepts employed by Group II is markedly different than those of Group I, the differences between the aims of the approaches that are adjacent to the watershed are not sharp in this regard, probably because most of the people involved were trained as psychologists. Nevertheless, the presence of a conceptual watershed is clear.

Social Judgment Theory

SJT has its origins in Brunswik's theory of perception (probabilistic functionalism), as may be seen from the prominent role played by such concepts as "ecological validity of cues," "utilization of cues," etc. It is worthwhile to note that probabilistic functionalism gave great emphasis to the question of the accuracy of perceptual judgments about, for example, the sizes of objects under various environmental conditions. Probabilistic functionalism thus had a strongly Darwinian tone because of its emphasis on achievement and adaptation.

In order to indicate the shift in subject matter from perception proper to quasi-rational judgment, interpersonal learning and conflict, that had become apparent in the 70s, as well as to indicate the introduction of new concepts, Hammond named this approach "Social Judgment Theory" (see Kaplan & Schwartz, 1975).

The probabilism that forms the major point of departure for SJT provides a conceptual link to the probabilism of all the approaches from economics. A major difference, however, is that whereas the probabilism of the approaches from economics requires explicit procedures for inquiring about a subject's probabilities regarding the occurrence of various conditions, events, or outcomes, SJT observes probabilistic behavior in those making judgments. Thus, while the SEU approach requires that the subjects directly "measure" their own uncertainty by, for example, choosing between lotteries, or other means, SJT measures the subjects' uncertainty (not their subjective probabilities) in terms of their

performance on judgment tasks, including interpersonal learning and interpersonal conflict. That difference grows out of the difference in the original aims of different approaches.

Information Integration Theory

IIT was developed by Norman Anderson in the 1960s. IIT has its origin in psychophysics; measurement of the subjective counterpart of physical stimuli is the point of departure for IIT. And whereas SJT denigrates the value of psychophysics as a subject matter for psychology, IIT sees the logic of psychophysical measurement as the essential foundation for the competence of psychological measurement, in more complex circumstances involving social circumstances as well as physical ones.

[MORE]

Attribution Theory

AT has its origins in the theoretical work of Fritz Heider who took Gestalt psychology as his point of departure (see *Thing and Medium*, 1926). Heider focused his attention on what he called "naive psychology" or "common-sense psychology," the cognitive strategies of persons who were required to make inferences or judgments about circumstances (e.g., causes) not immediately given in the midst of the ambiguities of life outside the laboratory.

Heider's book (1958) was highly influential and gave rise to several empirical and theoretical efforts to explore its utility. The most prominent and effective of these is AT.

Among the most frequently cited papers are those by Jones and Davis (1965), and Kelley (1967, 1973). These authors, among others, laid the groundwork for what has become perhaps the most prominent theory in social psychology during the 1970s. And it has drawn a large number of participants who have carried out a large number of empirical studies. We include AT in our attempt at integration of judgment and decision theory because AT is also a theory of judgment--perhaps the most psychological in character of all the theories to be discussed.

Integration

Group I approaches. DT, BDT, and PDT clearly have a common origin--the economic theory of choice. DT has pursued the logical, mathematical aspects of those theories, BDT and PDT have pursued the behavioral and psychological aspects. Prior to 1954, there does not seem to have been much intellectual contact between economics and the psychological approach to this subject matter. As Edwards observed at that time: the "voluminous. . . literature" in economics regarding "the theory of consumer's choice. . . is almost unknown to psychologists in spite of sporadic pleas in both psychological. . . and economic. . . literature for greater communication between the disciplines" (Edwards & Tversky, p. 13). Edwards further noted that he "could not find any thorough bibliography on the theory of choice in the economic literature" and as a result he included "a rather extensive bibliography of the literature since 1930" (Edwards & Tversky, p. 13). He did not mention that there was no "thorough bibliography" of the

psychological literature regarding probabilistic approaches to judgment and decision in psychology in 1954. Brunswik's 1952 monograph on the Conceptual Framework of Psychology, which gave some theoretical attention to probabilism, was known to him, but he did not refer to it, an omission that was symptomatic of the intellectual gap between Group I and II approaches. Edwards' main effort was in a different direction, namely, to construct the bases for a conceptual and empirical link between work in decision making in economics and in psychology. Which he did.

Edwards' familiarity with quantitative work in psychology, no doubt contributed to the fact that his 1954 and 1961 articles refer heavily to work that has strong roots in psychophysical measurement. At the risk of grand overgeneralization, the effort to link the theory of choice from economics with the quantitative experimental-empirical approach of psychologists can be described as an attempt to integrate utilitarianism with intuitionism under circumstances where both are given analytical (mathematical) and empirical (experimental) treatment.

[MORE]

Group II approaches. The origins of two of these approaches (SJT and AT) are close. Both Brunswik and Heider were Europeans who departed from the Gestalt point of view; Heider most notably in 1926 with the publication of "Ding und Medium," Brunswik in 1934 with "Wahrnehmung und Gegenstandswelt." Both could also be said to have reacted against conventional psychology, especially toward traditional psychophysics. Both could thus be said to have

an "ecological" approach in contrast to a "physical" one. (Note: Brunswik and Konrad Lorenz were contemporary graduate students (and friends) in Karl Bühler's laboratory in Vienna; it is not unlikely that both were attracted to a biological point of view as a result of Bühler's influence: personal communication from Lorenz to K. R. Hammond). But Brunswik's training as an engineer probably turned him toward quantification and empirical work, whereas Heider's more philosophical attitudes moved him toward more theoretical work. But the two arrived in the United States about the same time and were close intellectual companions who referred to one another's work frequently and with enthusiasm, both in print and in their lectures.

Hammond's "Psychology of Egon Brunswik" (1966) includes a chapter from Heider's book which illustrates the similarity in outlook, and a recent "Introduction to Attribution Processes" by Shaver (1975) includes a slightly abridged version of Brunswik's well-known lens model to explain the nature of Heider's approach to social perception. The interested reader should turn to Brunswik's (1952) development of "The Conceptual Framework of Psychology" for a description of the development of both approaches in the history of psychology, particularly with regard to their departure from both Gestalt psychology and psychophysics. And it should be noted that in 1958 Heider refers several times to the fact that Brunswik was making arguments similar to his as early as 1934 (see p. 56, for example). In fact, Heider's first reference to attribution, on whether failure will be "attributed" to

the person or the environment (p. 56) includes a reference to Brunswik (p. 220, 1934). In short, integration of AT and probabilistic functionalism was established in 1958, although that integration has virtually disappeared.

The one point that separated these two original thinkers during their lifetimes and still separates their intellectual offspring is probabilism (cf. Fischhoff); Brunswik enthusiastically incorporated it (in fact, was the first psychologist to do so), Heider minimized it, and one may guess (from personal communication: K. R. Hammond) that he found it distasteful. (See Holton, 1973, for a detailed account of the role of personal taste in the history of science; see also Brunswik and Krech, 1955, for a frank discussion of preference for scientific style in psychology.)

IIT is an off-shoot of the psychophysical tradition which Brunswik and Heider rejected. Anderson's interest in psychophysical matters has always been strong, but it is clearly different from Edwards'. Anderson saw the possibility not only of expanding the psychophysical approach to a multi-dimensional psychophysics (already underway by the 1950s) but also of extending it to social "stimuli." Which he did. This step is particularly evident in his applications of "functional measurement" to "impression formation" (from Asch) as well as to a wide variety of substantive problems of interest to those working in quantitative psychology as well as those in social psychology. Be that as it may, there is no doubt that the origins of IIT are very different from those of SJT and AT, and attempts to integrate their present work

will be superficial unless the fundamental metatheoretical positions taken by these approaches are explicitly discarded or somehow reconciled.

[MORE]

Since historians seem to seize upon clusters of events, with all due modesty, we offer them a hint by pointing to the following:

- 1954: Edwards' Theory of Decision Making article
- 1954: Thrall, Coombs and Davis' book on Decision Processes
- 1955: Hammond's application of Brunswik's probabilistic functionalism (of perception) to clinical judgment
- 1956: Publication of Brunswik's major work
- 1958: Publication of Heider's major work

In short, probabilism became a respectable idea, and judgment and decision making became respectable fields of study in the 1950s, and have remained so.

The independent and interdependent lines of developments will provide a rich source for the historians of social science, and the results of organizing the material should be valuable for understanding how the field of judgment and decision making became a cumulative scientific discipline. As a start, we offer this hint: Figure II-B-1.

[MORE]

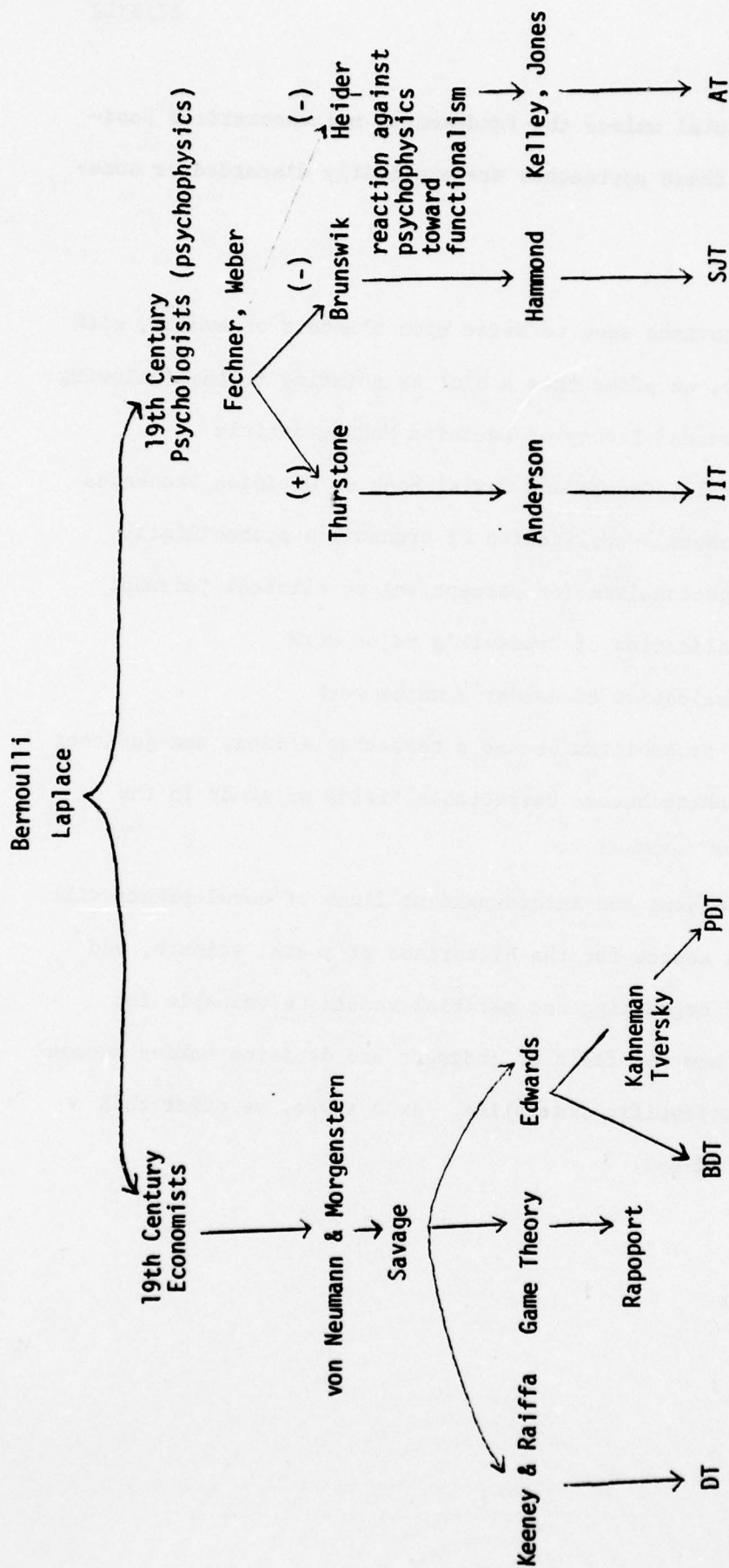


Figure II-B-1

C. Scope

Introduction

Judgment and decision theorists have cut out a large domain of behavior to be studied; this is particularly true if and when they recommend that the results of their research should be used. Indeed, Fischhoff (1976) has suggested that judgment and decision researchers stand in much the same scientist/practitioner role that clinical psychologists do. It is, therefore, highly important for judgment and decision making researchers to specify the scope of their theories, their empirical work, and their applications. That means disavowing competence in certain areas as well as emphasizing it in others. It also means acknowledging the fact that research tends to produce fragments of information, and, therefore, it will be difficult if not impossible for the judgment and decision analyst to bring to bear all the bits and pieces of knowledge that have been accumulated from piece-meal studies of part-processes to bear on a specific problem involving application. Worse still, not only will such leaps be difficult, they will be necessary, since applications will demand that it be done. As a result, efforts at application will always run the risk of involving unsupported generalizations across gaps in our knowledge. It is, in short, essential that the scope of various theories be somehow described, even though it is not in the nature of theorists to delimit and to disavow competence.

We shall use three broad descriptive categories in order to indicate differences in scope among different approaches, namely, the single-system case, the double-system case and the n-system case (see Figures II-C-1, II-C-2, and II-C-3). Again we refer to diagrams to indicate the meaning of these categories:

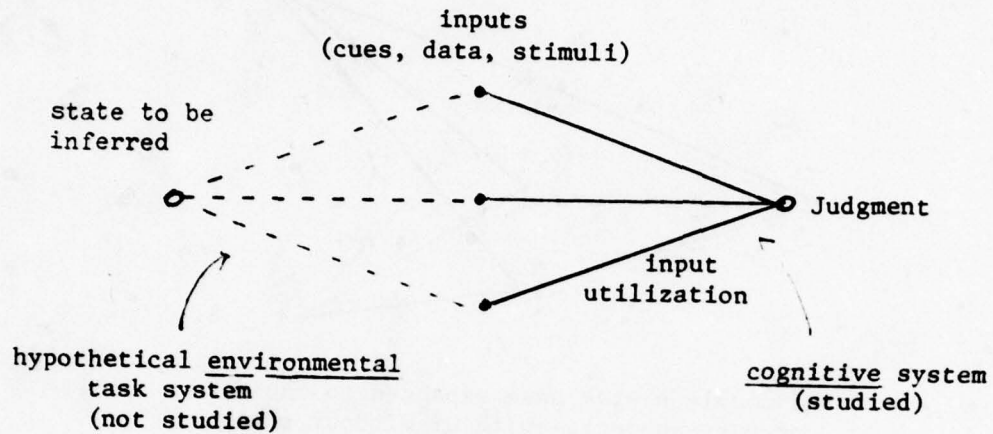
Decision Theory

DT restricts its theoretical interest to the single-system case, which involves one person making decisions without full knowledge of the task situation and without feedback about the effect of the decision. Since decision theorists do not theorize (systematically) about psychological processes or states, the scope of DT therefore is limited to those circumstances in which one person exercises his/her rational powers to the utmost under the guidance provided by a specialist in DT. But no analytical effort is devoted to the relation between the specialist and the client although it is often described in the language of popular psychology (see Fischhoff's analogy with psychotherapy for more).

[MORE regarding hierarchical models]

Since both DT and game theory follow from von Neumann and Morgenstern's classic work, game theory can be seen as an approach to the n-system case with a parallel development to DT. Keeney and Raiffa make little use of game theory of the Luce and Raiffa (1957) variety or that pursued by Anatol Rapoport (1974) with his special emphasis on the Prisoner's Dilemma game that, strangely, captured the interest of psychologists for so long. And although

Single-system Case



(Dashed line indicates no data collected.)

Figure II-C-1

Double-System Case

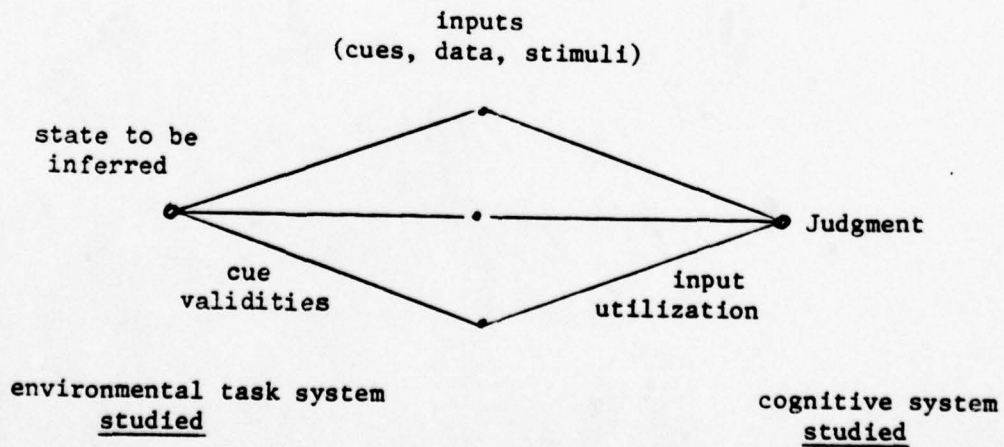
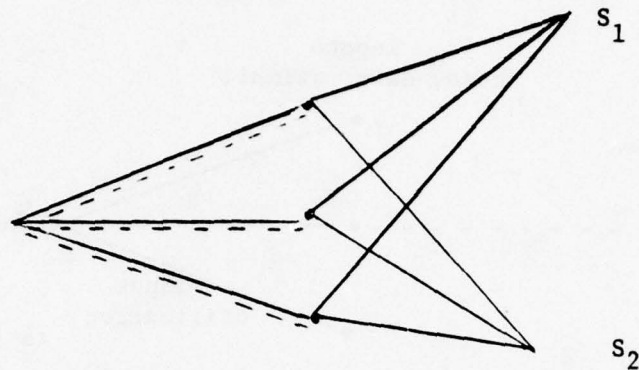


Figure II-C-2

N-System Case



the double-system case expanded to include
2 or more subjects--with or without study
of the environmental (task) system

Clearly, the n-system case can be broken down into various
types, but for our present purposes these distinctions are enough.

Figure II-C-3

Keeney and Raiffa indicate an interest in the problem, they treat the matter of interaction largely on a "clinical" level.

DT work on social welfare functions and group decision making (Chapter 10 of Keeney & Raiffa, 1976) might also be viewed as studies of the n-system case. However, the basic approach to such problems is to reduce the group decision to an individual decision of a supra decision maker (see also, Harsanyi, 1974); in effect, the n-system case is converted into the single-system case so that existing DT can be directly applied to the group situation.

An interesting extension of the SEU model has recently been provided by McClintock (1977) who applies the model to the n-system case and studies interpersonal conflict [and introduces new concepts]. This extension is discussed below in connection with BDT.

Behavioral Decision Theory

Because BDT is directed toward reconstruction of SEU theory in light of the experimental analysis of decision and choice behavior, BDT theorists found it necessary and useful to do experiments. The primary reason for doing experiments was to increase our understanding of how persons process information and to discover the impact of information on the revision of subjective probabilities. Thus external conditions were manipulated in order to test hypotheses regarding internal conditions, and thus objective (external) probabilities were manipulated, first to test the extent to which the axioms of SEU theory could account for the behavior of decision makers, and then to examine the extent to

which the Bayesian model could account for revision of subjective probabilities. Manipulations of objective probabilities (including conditional probabilities) could thus lead to observing, for example, the impact of a datum (information) on the revision of a probability of an event. In order to do this, the investigators had to construct task systems (the properties of which would be fully known to the investigators). Once the tasks were constructed, comparisons could be made between the properties of various decision tasks (e.g., their unconditional probabilities and conditional probabilities) and the properties of the cognitive (decision) systems brought to bear on the tasks (e.g., the parallel probabilities). BDT theorists could thus examine the relations between external and internal unconditional probabilities, describe the differences between them, and explain why such differences occurred.

Probabilities are again treated as judgments, as in the case of DT, but they are now related to objective probabilities and the relation between the two is treated as if it were a scaling problem in psychophysics. The theoretical examination of the relation between the two includes concepts taken from psychophysics, as for example, anchoring, etc. In addition, BDT theorists found it useful to compare the conditional probabilities in the cognitive system of the decision-maker with those in the task. The conditional probabilities typically concern the probability of a distal event given the occurrence of some proximal event. But inter-event conditionality is considered as well.

Research that included the presence of, and documentation of the properties of, various decision tasks meant that BDT theorists not only were interested in the single-system case but the double-system case as well (Hammond, Stewart, Brehmer, Steinmann, 1975, erroneously described BDT as strictly single-system). The necessity for the shift from the single-system case to the double-system case marks an important distinction between the scope of the theoretical effort of DT and BDT, and symbolizes as well as diagrams the distinction between the locus of the concepts employed by each theory. (This distinction is also illustrated by Shanteau's diagram; see Figure II-E-3.)

Introducing an independently real second system into the analysis requires BDT theorists to be specific about the properties of the decision task system. That requirement meant that some set of terms must be employed to describe the task. The terms chosen by the BDT theorist are statistical terms (unconditional and conditional probabilities, states of the system, dependencies, etc.). Note that the terms used by BDT to describe cognitive activity are parallel to the terms used to describe the environmental task.

[MORE; especially on cascaded inference]

BDT has not, however, extended its theoretical or empirical hypothesis-testing research efforts to include interpersonal conflict (or learning) with regard to decisions although it has extended its application to that situation (the n-system case).

Psychological Decision Theory

This approach was developed in the context of the single-system case, and some extensions have been made to the double-system case (see for example, the work on "illusory correlation"). But generally, a comparison is made between (a) inferences justified by statistical logic and (b) inferences made by persons from the same data, in contrast to tasks involving independently real events. So far no work has been done with regard to the n-system case.

Transition

Moving from Group I approaches to Group II approaches enlarges the scope of the behavior covered by both theory and research. For whereas the Group I approaches considered here intensively focused on measurement problems related to subject probabilities, utilities, and their combinations with respect to such matters as transitivity of choices, etc., the conceptual framework that applies to more than one decider bears an uncertain relation to DT, BDT, and PDT. (The substantial amount of theory and research that flows from game theory and social welfare function theory, however, needs eventually to be considered in detail here, for it considerably enlarges the scope of the work within Group I approaches.)

Group II theories have produced a substantial amount of research involving two or more persons involved in judgments, decisions, and inferences. These approaches have, to various degrees, also developed theoretical concepts intended to apply to the interaction between persons. As we shall see, SJT has extended

the original lens model to include cases involving two or more persons, IIT has inspired at least a few studies involving groups, and AT has been applied to a wide variety of individual and group circumstances.

Social Judgment Theory

This approach includes within its scope the single-system case as well as all other cases and Brehmer in particular has carried out a large number of multiple-cue probability studies (double-system case) and a large number of studies on interpersonal learning and interpersonal conflict (n-system case). These studies are fundamental insofar as SJT is concerned and provide the basis for generalization to the analysis of the single-system case. Thus SJT intends to include within its scope the following topics: (a) learning under uncertainty (MCPL), (b) interpersonal learning, (c) interpersonal conflict, and (d) group judgments and decisions. The work that has been done in these areas is intended to provide the basis for conflict resolution in situations involving policy disputes.

[MORE; especially on hierarchical models]

Information Integration Theory

This approach was developed to deal with the single-system case, although it has been applied to both the double-system case (involving learning) as well as the n-system case (involving group judgments). There is no reason in principle why IIT should not

be or cannot be applied to the study of interpersonal learning or interpersonal conflict.

Attribution Theory

This approach is highly general in its aim. It covers not only the case of a single person making judgments about causal locus, but interpersonal relations between two or more people. Its intended scope is large and includes references to interpersonal learning and conflict. No boundaries are implied in AT.

Integration

Group I. Proponents of the three approaches within Group I generally have not, in general, widened the scope of their theoretical endeavors beyond the single-system case. DT gives virtually no attention to empirical research carried out within the double-system case by BDT's such as Edwards and his colleagues other than to provide footnote acknowledgements of the existence of such work (see p. 8 and p. 212 of Keeney & Raiffa, 1976). And although there is mathematical work within the general framework of SEU theory on coalitions and interpersonal comparisons of utilities, it does not seem to appear relevant to DT as described by Keeney and Raiffa, 1976. (Neither topic is indexed.) Nor is there any significant reference in Keeney and Raiffa to work on conflict resolution by theorists such as Anatol Rapoport or to the psychological research on interpersonal conflict by SJT, although work by Luce and Raiffa, Arrow, and Harsanyi is mentioned. One might guess that work by DT in this area will be more likely

to extend in the direction of welfare functions than game theory. But what if decision theorists like Keeney and Raiffa (or Zeleny) become more clinical? Psychologists are certain to find the remarks of Keeney and Raiffa (and DT generally) regarding conflict resolution to be hopelessly naive, however sophisticated they may be from a logical viewpoint. Keeney and Raiffa do refer to the use of sensitivity analysis as a conflict-reduction technique, but that is about as far as they go. And although Edwards observed in 1961 that the early work on group decisions "had a substantial resemblance to the SEU model," there has been no systematic theoretical unification of individual decision making and conflict behavior by DT. And with the increasing challenge to the basic premises of the traditional SEU model from PDT and other approaches (e.g., IIT), it is doubtful that the extension of SEU theory to group decision making is likely to be accepted without also being challenged by the other approaches.

There are psychologists within the BDT approach who are now attempting to extend the SEU approach to the n-system case and thus to study interpersonal conflict. McClintock's recent article in Druckman (1977) provides an excellent example.

Proposition 5. In situations of social interdependence, one's access to one's own valued outcomes is dependent upon the values assigned by others to their possible outcomes (and vice versa). Furthermore, the attractiveness of outcomes for one actor may be influenced not only by the outcomes he receives but also by the outcomes he judges the other(s) will receive.

Group II approaches. The three approaches grouped here appear to be highly similar in intended scope. SJT has been explicit in its efforts to study all three cases mentioned above, and thus intends to contribute to the literature regarding judgments, learning regarding a physical environment, learning in a social environment (e.g., interpersonal learning from and about another person) and interpersonal conflict; research in all of these areas applies the same systematic framework; work on the n-system case has continued since about 1965; work with other cases began earlier. But there has been no attempt to relate this work to IIT or AT or incorporate any of the concepts from either into SJT research. (But see Mumpower, 1976, for a detailed analysis of the similarities and differences between SJT and AT.) The same is true for IIT and AT with regard to SJT.

There have been attempts by Anderson (1974) to integrate IIT and AT within the framework of the single-system case, and to some extent, the n-system case as well. The reverse is also true; Jones, Kanouse, Kelley, Nisbett, Valins and Weiner (1972) make use of certain IIT concepts and measurement techniques in (primarily) the single-system case. There is no reason, in principle, why the techniques employed by SJT to study interpersonal conflict could not be employed by IIT. The effects of different subjective scale values and/or weights could be studied in exactly the same way. Most important, the approach could readily examine the differential effect of various integrative mechanisms (adding, averaging, multiplying, etc.) on interpersonal conflict just as, for example,

Brehmer and Hammond (1973) examined the differential effect of the use of various function forms in interpersonal conflict, and as Brehmer has examined similar effects on interpersonal learning. This is perhaps the easiest situation for integration and potential cumulative information we shall encounter. (In addition, interest in the n-system case could lead to the evaluation of the role of equal weights [see methodology section] in conflict reduction.)

The single-system case: producing cohesion or confusion? In this situation there is no environmental task "out there," the characteristics of which are independently known. Do you choose to merge your company with this one or that one? Choose one or the other course of action and you will never discover what the outcome of the action not chosen would have been. In this case, the "right" choice will never be empirically known because taking one action ordinarily precludes the possibility of discovering what a different action might have accomplished. (Einhorn has considered the psychological implications of this situation in detail.)

(Contrast the single-system case with the double-system case, in which the right answer can be ascertained more or less readily; predicting the weather or the activity of the stock market or a horse race provides examples. Cases involving interpersonal conflict and interpersonal learning illustrate more complex situations that have come under study by judgment researchers in which the right answer also can be empirically determined.) But, all six approaches share a common interest in the single-system case. Indeed, it is because

of the shared interest in this specific case that this attempt at integration is being made; this case provides the link between Group I and II as well as within Groups. And it is here that confusion is likely to arise over the intended scope (and function) of various approaches. Therefore, at the risk of repetition, we offer what appear to us to be important distinctions in scope (and aim) with regard to this case.

1. DT and BDT offer descriptions not of what is taking place but rather of what ought to be taking place and offer assistance in reducing the gap between the two. DT and BDT declare that their intentions are to help their clients "get their heads (or preferences) straight," or something similar. Careful delineation of probabilities and utilities aggregated by the decision analyst provide the remedy (see, Keeney, 1977, for a detailed, worked out example). Although probability assessments are treated with some sophistication and steps are often taken to straighten out logical flaws, no specific remedies other than those based on "clinical" experience are offered. (This is particularly true with regard to interpersonal dispute over probabilities and utilities, in regard to which discussion (!) is offered as remedy.)

2. PDT is primarily interested in explaining why departure from rational choice occurs in situations where probability assessment is pivotal, and on the basis of these explanations offering remedies (debiasing techniques) for the biases observed. Note that PDT offers remedies based directly on its empirical research. Thus it offers specific debiasing remedies for the specific biases it discovers in the single-system case. PDT, does not however,

address the double-system case in which answers are empirically produced. (Answers are produced through calculations based on statistical logic.) Nor does PDT address the n-system case, although it would not find it difficult to do so. Indeed, its psychological orientation and concepts (availability, etc.) offer strong suggestions as to the source of difficulties in interpersonal learning and interpersonal conflict. Thus, the potential scope of PDT does include the n-system case. And if one recalls the work in the area already carried out by SJT and the similar potential for work by IIT, a strong case could be made for the argument that such work should be pursued because of the complementary character of that work. This opportunity for cumulativeness should not be overlooked.

Domains of behavior of theoretical interest. There is one further question that can be addressed to the intended scope of various approaches: insofar as they intend to be descriptive or explanatory, what is it that they intend not to describe and/or explain? Do these approaches intend to limit their descriptions and/or explanations in any way? Although these approaches seem to be limited to certain domains of behavior and in the breadth of their intentions, they are not specific about their boundaries. For example, with the exception of AT none of these approaches include concepts that apply to domains of behavior other than cognition (although there is certainly room for Group I approaches to claim an interest in motivation). No approach includes theoretical concepts that apply to states other than normal ones.

Thus, for example, stress, affect, specific need-oriented states are generally ignored, as are other processes (e.g., memory; but see Hamilos & Pitz's, 1977, research on memory and judgment; see also Tversky & Kahneman for references to memory in their work). Gillis (1975) and Hammond and Joyce (1975) have shown how SJT can be profitably extended to the study of the effects of psychoactive drugs on judgment in the single-system, double-system and n-system cases in populations of psychotic patients. But limits need to be drawn, if only tentatively.

D. Intended FunctionIntroduction

What are current theories of judgment and decision making supposed to do? What are their intended functions? Prescribe? Describe? Explain? Predict? Without knowledge of their intended function, theories can hardly be compared. Unfortunately, judgment and decision theorists have not always made their theoretical aims entirely explicit. And they are also apt to be vague with regard to the limits of competence of their theories. What sort of competence do they disavow? What sort of agreement exists among the six approaches with regard to intended function? Is everyone trying to do the same thing? Or to complement the work others are doing?

Decision Theory

This approach focuses on decision-making from a prescriptive point of view only (see Keeney & Raiffa, 1976, introduction). Its prescriptions are based on the utility theory developed by the mathematician von Neumann and the economists Morgenstern and Arrow, among others. In its recent development, Keeney and Raiffa (1976) emphasize the point that the aim of DT is to elaborate the logical entailments of subjective probability and utility theory and extend them to a variety of circumstances by means of mathematics. The criterion for the validity of the theory is its logical, mathematical consistency. Once developed, the theory stands as a

logical structure for decision making; decision makers may then use it in order to achieve the logical consistency provided by the theory. The achievement of logical consistency is, therefore, the reason for the use of DT in the real world.

DT makes no claim that it represents or describes the cognitive activity (or information processing) of human decision makers. Indeed, it is precisely because of the presumed departure of decision making from the logic of decision theory that causes decision theorists such as Keeney and Raiffa to insist that people, especially policy makers, should change their decision making behavior to make it conform with the precepts of DT. Accusations, therefore, that DT does not represent the cognitive activity of any person does not deter decision theorists from developing new applications or pursuing the implications of a theorem. The emphasis is not on what decision makers do, but what they should do. Explanations of why human decision makers deviate from the logic of DT is a matter left to psychologists.

Although that point has been made many times, we repeat it here because current criticism by psychologists of SEU theory is directed precisely at its intended function, not its logical structure. Thus, it is argued that however satisfying SEU theory might be from a mathematical point of view, it is not useful as a guide to decision making because human beings do not behave in accordance with the fundamental assumptions of the theory. When, for example, empirical evidence indicates that the premises of SEU theory do not represent actual choice behavior, the validity

of the postulates of the SEU theory is denied, and thus the validity of the behavioral entailments of the theory is denied.

One might expect decision theorists to find such criticism to be misdirected on the ground that it is based on a faulty interpretation of the intended function of the theory. DT argues that (a) if one must decide or choose, then an additive or multiplicative combination of expectations and utility is an appropriate basis for decision or choice and (b) the logic of DT, as articulated by mathematical analysis, provides the best guide for reaching defensible decisions. It is irrelevant, therefore, so far as DT is concerned whether (unguided) behavior is in accord with the axioms of SEU theory. What is important is that decision making behavior should be in accord with these axioms. They further believe that once the axioms are carefully explained, any reasonably intelligent person would want to change his behavior to be in accord with these axioms. As a result, if and when psychologists (or others) find behavioral violations of the axioms, decision theorists dismiss such discoveries as irrelevant to their purposes. Thus, for example, Raiffa (1968, p. 77) states:

"The following example will help to illustrate how intransitivities may arise in descriptive choice behavior and why in a prescriptive theory of choice this type of behavior should be discouraged."

In short, DT considers its axioms to be reasonable and desirable rules for decision making behavior that everyone would want to follow, once they understood them. DT thus disavows any intention to provide explanations for why decision

making takes the form others claim it does, for it has only prescriptions for improving it. And that, they appear to be saying, is enough. As a result, empirical challenges to the basic postulates of DT have been ignored by Keeney and Raiffa, and management scientists in general.

But Cochrane and Zeleny (1973) are management scientists (or operations researchers) who do not share this aloofness, and as we shall see below, psychologists are in sharp disagreement with it.

Behavioral Decision Theory

This approach intends (a) to describe departures from optimal decision-making in empirical detail (in contrast to DT) and (b) it also intends to explain such departures in terms of both external (task) and internal (psychological) conditions. Perhaps the best-known description of a departure from optimal decision-making is conservatism. This term signifies the putative failure of decision makers to revise their posterior subjective probabilities as much as they should (as determined by Bayes' rule for aggregation of conditional probabilities in relation to hypotheses) upon the receipt of new information, a (disputed) finding that led BDT theorists to describe decision makers as "Conservative Bayesians."

Edwards made it clear in his landmark article of 1954 that he considered the economists' theory of value and choice to be a valuable guide for the empirical work of psychologists, and that psychologists should carry out empirical tests of the implications of the theory rather than dispute its assumptions (Edwards &

Tversky, 1967, p. 16). By 1967 Edwards was still of this persuasion: "All these topics represent a new field for psychologists in which a theoretical structure has already been elaborately worked out (by economists: K. Hammond), and in which many experiments need to be performed" (1967, p. 55). That is, psychologists should discover to what extent deductions from theories of choice and preference are empirically true. Therefore, the intended function of BDT is to reconstruct, if, as, and when need be, the already worked out theory of decision or choice behavior, on the apparent assumption that such theories intend to be and should be a true description of the behavior of persons making decisions.

Such a confirmation or reconstruction of theories of choice would place a powerful prescription for decision making in the hands of decision analysts who could then provide "tools" for decision makers. The culmination of the effort can be seen in Edwards' paper presented during the 1977 meetings of the American Association for the Advancement of Science at a symposium entitled "Judgment and Decision in Public Policy Formation" ("Technology for Director Dubious: Evaluation and Decision in Public Contexts," 1978). Here Edwards refers to ". . .two relatively simple tools, both currently in use, in contexts in which they obviously bear on public policy, and could be used by public policy makers, but so far have not been" (p. 73). (A surprising statement in view of Keeney & Raiffa's work with these same "tools.")

It now appears that Edwards and his colleagues intend to pursue the study of choice behavior through empirical experimental

analysis within the Bayesian framework, and to pursue application of their findings in conjunction with SMART--a simplified version of MAUT.

[MORE]

Psychological Decision Theory

PDT intends not only to describe but to explain the discrepancies between human decision behavior and the fundamental premises upon which SEU DT is based. Indeed, the proponents of PDT have made it clear that they believe that they, among others, have already shown that the basic premises of SEU theory of decision behavior are empirically false, and, therefore, that a different description of human decision behavior should replace it (see Kahneman and Tversky's, 1977, Prospect Theory paper, also Slovic, Fischhoff & Lichtenstein, 1977). The proponents of the PDT approach have carried out empirical research that, in their view, and in the view of many others, not only shows that people do not make decisions and choices in conformity with the precepts that support DT and BDT, but explains why people do not. This research has been shown to be a fruitful combination of theory and empirical work in that new (or almost new) theoretical concepts have been introduced and their explanatory value tested empirically.

The approach of PDT is the one psychologists would expect:

(a) identify some important behavioral phenomenon (e.g., departure from optimal choice), (b) adduce some psychological concepts (e.g., availability, representativeness) to account for or explain why the

discrepancy should occur, and (c) conduct experiments to test the validity of the explanatory concepts. Thus, it might be said that if (a) the intended function of BDT is to shift the study of decision behavior from a logical path of research (that led to SEU theory) to a behavioral path, then the intended function of PDT is to move us from a sheerly behavioral path to a psychological path that explains why decision makers behave the way they do. Have we now reached a different point from that described by Edwards when he indicated that "a theoretical structure has already been worked out"?

Not only does PDT offer to explain why "people replace the laws of chance by heuristics" they are now prepared to indicate when people are apt to make this replacement. Although Slovic, Fischhoff, and Lichtenstein observed in 1977 that "heuristics (PDT) may be faulted as a general theory of judgment because of the difficulty of knowing which will be applied in any particular instance" a recent paper (Causal Schemata in Judgments under Uncertainty, Tversky & Kahneman, 1977) takes steps toward removing that "fault," thus: "Distributional data affect predictions when (*italic ours*) they induce a causal model which i) explains the base rate, and ii) applies to the individual case." And conditional statements are made, thus: "Distributional information which is not incorporated into a causal schema, either because it is not interpretable as an indication of propensity or because it conflicts with an established schema, is given little or no weight in the presence of singular data."

Certain predictions can also be found in earlier papers. Thus, in 1972; "the representativeness heuristic is more likely to be employed when events are characterized in terms of their general properties; whereas the availability heuristic is more likely to be employed when events are more naturally thought of in terms of specific occurrences" (Kahneman & Tversky, 1972, p. 452). "When the generic features of an event as well as its specific features are considered, both heuristics are likely to enter into the evaluation" (Kahneman & Tversky, 1972, p. 452; italics ours). In short, although PDT has not been explicit (so far as we know) about whether its intended aim is that of prediction as well as explanation, PDT is clearly well on the way toward that goal.

The clear aim of PDT to enter upon the enterprise of providing psychological explanations of probability (and utility) judgments has led to the sharp challenge of the validity and thus utility of DT (a challenge that has been ignored, as we mentioned earlier). In short, PDT argues that the logical path taken by DT, whatever its value in other fields of operations research, does not lead to a defensible prescriptive theory of decision making behavior.

It is important to note that PDT is rooted in decision theory and that it still includes probabilities and utilities among its central descriptive terms, and that one of its intended functions is to evaluate the decision making behavior of subjects in terms of optimality prescribed by a mathematical (statistical) model.

As noted above, evaluation largely takes place with regard to the manner in which subjects assign probabilities to events. Thus,

PDT intends to discover whether biases occur in this type of cognitive activity. Since PDT claims that biases are highly prevalent, it offers suggestions as to how specific biases might be removed, an aim that has received the name of debiasing. Debiasing, it should be noted, is a far different type of assistance than that offered by DT and BDT.

The intended function of PDT with regard to aiding the decision maker is perhaps best summarized by Slovic (1976, p. 238) as follows:

Subjective judgments of probability and utility are essential inputs to decision analysis. We still do not know the best ways to elicit these judgments. Now that we understand many of the biases to which judgments are susceptible, we need to develop debiasing techniques to minimize their destructive effects.

The complexity of that task is also indicated, however;

Simply warning a judge about a bias may prove ineffective. Like perceptual illusions, many biases do not disappear upon being identified.

And, indeed, Lichtenstein, Slovic, Fischhoff, Layman and Combs (1978), have shown that debiasing is not a simple matter, and Kahneman and Tversky's own work (1978) has not offered empirical proof of the efficacy of their debiasing efforts. Shanteau and Phelps, however (1977) show that some debiasing techniques may be carried out effectively, albeit unwittingly, by laymen.

Transition

As noted in the sections on origins, moving from the PDT approach to SJT marks a move from three theories that have their origins in economics to three theories that have their origins in psychology. (Note: The academic origins of theories are not to be

confused with the academic origins of the authors.) And although the first psychological approach (SJT) to be encountered on this side of the conceptual watershed includes probabilism in both environment and organism as a basic premise, which, of course, links it to SEU theories, neither it nor the other two psychologically-oriented theories include either probability or utility as basic descriptive terms or as basic units of analysis that require measurement. Neither probability nor utility are, as Slovic put it with regard to decision theory, "essential inputs to decision analysis," once the conceptual watershed is crossed. As might be expected, therefore, the intended function of the three theories to be discussed next is different from the previous three. Just how different is unclear and the reader should anticipate transitions rather than abrupt changes in the conceptual terrain.

Generally, the three psychological theories are theories of (quasi-rational) inductive knowing; that is, they are directed toward the question of how human beings acquire or apply knowledge under circumstances of ambiguity in the task. Ambiguity is always present in the studies of problems that Group II theorists study, for it is ambiguity that creates the knowledge problem. Ambiguity, however, is described in different ways by different theorists. Indeed, as we shall see, identification and description of the sources of ambiguity is one main topic that differentiates these theories. Group I approaches, on the other hand are generally indifferent to analyzing, or theorizing about, the sources of ambiguity in environmental tasks (although PDT frequently points to them).

Social Judgment Theory

The intended function of this theory is to describe, but not explain, human judgment processes. In its formative stages (1955-1970) of applying Brunswikian concepts to judgment tasks its efforts were directed toward social perception (see e.g., Hammond & Kern, 1959; Crow & Hammond, 1957). Since 1970 its efforts have been directed less toward establishing the accuracy of its descriptions than toward the usefulness of its descriptions for the policy maker, within, of course, certain boundaries of accuracy (see Hammond, Stewart, Brehmer, & Steinmann, 1975). It should be noted, however, that at least one SJT researcher (Brehmer) does direct his efforts toward explanations, both in the case of multiple cue probability learning as well as interpersonal learning and interpersonal conflict, and offers and tests hypotheses that are explanatory in nature (see e.g., Brehmer, 1975). In addition, SJT intends to provide guides toward the development of judgment (and decision) aids.

There is no such thing as description without theory, of course, and therefore there is an explicit theoretical as well as metatheoretical basis for the descriptions provided by SJT.

The strong emphasis placed on the problem of knowing (in contrast to choice) and the explicit theoretical attention given to the environment can be seen in the following quotation from Hammond, Stewart, Brehmer, and Steinmann (1975):

WHY IS JUDGMENT REQUIRED?

Knowledge of the environment is difficult to acquire because of causal ambiguity--because of the probabilistic, entangled relations among environmental variables. Tolman and Brunswik called attention to the critical role of causal ambiguity in their article "The Organism and the Causal Texture of the Environment" (1935), in which they emphasized the fact that the organism in its normal intercourse with its environment must cope with *numerous, interdependent, multiformal relations* among variables which are *partly relevant and partly irrelevant* to its purpose, which carry only a *limited amount of dependability*, and which are *organized in a variety of ways*. The problem for the organism, therefore, is to know its environment under these complex circumstances. In the effort to do so, the organism brings a variety of processes (generally labeled *cognitive*), such as perception, learning, and thinking, to bear on the problem of reducing causal ambiguity. As a part of this effort, human beings often attempt to manipulate variables (by experiments, for example) and sometimes succeed--in such a manner as to eliminate ambiguity. But when the variables in question *cannot* be manipulated, human beings must use their cognitive resources unaided by manipulation or experiment. They must do the best they can by passive rather than active means to arrive at a conclusion regarding a state of affairs clouded by causal ambiguity. They must, in short, exercise their judgment. Human judgment is a cognitive activity of last resort.

It may seem odd to remind the readers of this volume of the circumstances which require human judgment, yet it is essential that we do so, for it is precisely these circumstances which are so often omitted from studies of human judgment. If we are to understand how human beings cope with judgment tasks, however, not only must such ambiguity be present in the conditions under which human judgment is studied, but causal ambiguity must itself be represented within the framework of a theory of human judgment (Brunswik, 1952, 1956; Hammond, 1955).

The absence of the terms subjective probability, utility, choice, and preference is apparent.

Information Integration Theory

IIT is a psychological theory that intends to discover (cognitive) psychological laws that intervene between stimulus and response and thus explain, or at least account for, the relation between S and R. IIT intends to describe human cognitive theory (of which decision, judgment, and attribution are merely special

cases) in quantitative terms and to account for such activity in terms of "cognitive algebra" (see Anderson, 1974, p. 84). More specifically, IIT focuses on the organization or integration of information by describing the adherence to (or departure from) various algebraic formulations such as additive equations, averaging equations, etc., that are treated as "models" of cognitive functioning. Therefore, a major part of the descriptive effort lies in discovering which model "best fits" the relation between S and R. Does a model that adds, or averages, or multiplies its terms "best fit," i.e., best describe, best account for, the cognitive activity of integrating information of any kind that takes place under specified circumstances? Which, of such models, generally "fail" as descriptive devices?

But if it is clear that IIT is a law-seeking endeavor, it is unclear what generality it seeks. It surely is not in the scale values assigned to the stimuli in any task-situation, nor is it in the weights which subjects might assign to specific stimuli, nor does it seem even to lie in the generality of the organizing principle (integration rule) discovered--since Anderson and others are prepared to find persons adding, averaging or multiplying virtually any set of weighted, scaled stimuli (unless a previous study has found one result or another). We do not find predictions of the appearance of any specific organizing principle other than on the grounds of a precedent. Rather, the search for generality seems to focus on a more general level. Thus, for example, in Berkowitz (1974, p. 84) Anderson says: "The present survey has

yielded considerable support for the operation of a general cognitive algebra in social judgment. Cognitive integration seems to follow simple averaging, subtracting and multiplying rules far more commonly than has been recognized" (*italics added*). (Was "adding" intentionally omitted?)

We draw the conclusion from these remarks (and others) that for IIT, explanation and mathematical description converge, and from that convergence emerges a psychological law--precisely as in physics. Thus, if $F = ma$ is a mathematical description of a relationship that can be applied to the behavior of a wide variety of objects under a wide variety of conditions, it is also explanation, and it is also physical law. And in parallel form, if a specific form of cognitive algebra is regularly found to describe with precision the cognitive activity of a wide variety of people under a wide variety of circumstances, then, just as in physics, that specific form of cognitive algebra is explanation enough. And if enough regularity is achieved over enough conditions, lawfulness is achieved.

Theory and a specific form of methodology are so tightly interwoven in IIT that it is difficult to separate them. One can hardly escape the conclusion that at least one intended function of IIT is methodological prescription, namely, to convince judgment and decision researchers on methodological grounds that the problem itself requires a certain method, namely, functional measurement.

Functional measurement is strongly related to psychophysical measurement. Indeed, the intended function of IIT appears to be

that of applying a new and more complex form of the psychophysical method of measurement to (somewhat) complex social, rather than physical, data. (The SJT or Brunswikian approach, for example, is dismissed as applying only to subjects' use of physical stimuli; see Anderson, 1974.) Many of the conditions examined by IIT for their effect on cognitive activity include those employed in traditional psychophysics (e.g., serial effects, anchoring, set size, contrast, etc.). And, of course, the large emphasis on stimulus scaling, response scaling, the concern with the correctness of the "Power Law," and the general respect for precision reflect the intended function of this approach to discover the laws of "information integration" that include not only traditional psychophysics but social data or stimuli as well. Therefore, unless (a) other approaches agree that they also are studying a more complex form of psychophysics and develop their work accordingly, or (b) IIT shows that irrespective of what proponents of other approaches may believe, their work is in fact only a variety of a more complex psychophysics, then IIT will very likely find its results to be nonacceptable to proponents of other approaches, despite strong efforts to incorporate other theories into IIT via the techniques of functional measurement.

Support for that general view can be found in the persistent didactic tone in Anderson's statements. Methodological admonitions can be found in virtually all of the writings of IIT theorists, and although the admonitions can only be admired from the point of view of theoretical-methodological integrity, they are apt to be

less than fully welcome from the point of view of a theorist with a different aim or conception of what the judgment and decision problem is, or with a different evaluation of the psychophysical method, multi-dimensional or not.

Admonitions that reflect the intended function of IIT can also be found with regard to theoretical style, e.g., "Those who work with models. . . find that they impose a far tighter conceptual discipline than the more verbal formulations. Indeed, the attractiveness of verbal theories sometimes seems to reflect the verbal glitter of surplus meaning in the theoretical terms rather than any implicational power in the theoretical structure. Balance theory is a good example of unfilled promise. . . and recent work on social attribution often seems to place undue weight on a clever theoretical story" (Anderson, 1974, pp. 85-86).

It should be noted that this same sort of didactic, admonitional tone can be found in Hammond's, and to some extent, Edwards' writings as well. We emphasize this tie between the advocacy of theory and method in Anderson's writings because of the clarity with which it is expressed.

Attribution Theory

The general aim of this approach is to describe and explain the cognitive activity of persons as they attempt to describe and explain the behavior of objects and events in their customary habitat. Thus, it is intended to be the psychology of "common sense." And therefore, the locus of its intended function is virtually the same as in the case of SJT.

More specific aims have been developed by Kelley and Jones in what has now become "attribution theory." They intend to explain why persons attribute the causes of other persons' behavior to forces internal or external to them. AT differs from the two previous theories in that it is largely indifferent to quantification of the data used as the basis of attribution, although this is not a formal position and there are some exceptions (mostly related to IIT or Bayesian work; see Fishbein & Ajzen, 1975, for extended example of the latter).

Despite AT's recent (c. 1970) emphasis on the attribution of internal vs. external causality, the general theory developed by Heider (1958) is indeed a general theory; it disavows little with regard to its explanatory domain, and very little, if anything, with regard to matters it considers to be worthy of the attention of those psychologists who are interested in behavior outside the laboratory. Because it shares much of its aims with SJT further discussion of the intended function of AT is deferred to the integration section.

Integration

Group I theories. Among these theories we find several aims: (a) to describe, (b) to explain choices of action and preferences for various states of the world, and (c) to provide decision aids for persons making such choices and preferences. Although different emphases are placed upon (a), (b), and (c) by different Group I approaches, it is not these differences in general aim that create

the sharp division among them. Rather, it is differences at a more specific level. For example, although DT is content to accept the notion that unaided decision makers are less than rational, BDT intends to describe departures from rationality. "Conservatism," of course, is the classical example. Note that BDT (Edwards) is strongly committed to the Bayesian model with regard to probability estimates in the SEU model and therefore intends to describe choice or decision behavior in terms of the degree or form of departures from the Bayesian model, rather than in terms of a model that fits the behavior itself.

PDT approach is not satisfied with that aim. It goes further in its explanatory effort than BDT in that it not only is concerned with departures from optimality, or rational choice, but is concerned with the psychological process that accounts for the subject's decision. Indeed, the results of PDT research challenge the validity of the premises upon which DT and Bayesian BDT research rests. And insofar as these challenges find empirical support they may well result in the development of an entirely new "decision theory," one that is based on empirical behavioral regularities discovered in the psychological laboratory. (See Slovic, Fischhoff, & Lichtenstein for a review of empirical studies that contest the validity of the ". . .staggering SEU model. . ."[1977, p. 11].)

In sum, there is a progression in intended function from (a) DT which concerns itself only with the "best process," (b) BDT which examines and describes departures from the "best process" in largely statistical terms, and (c) PDT which not only explains departures,

but substitutes new descriptions, models or theories of how subjects make decisions in place of the rational (SEU or Bayesian) model from which the subjects' decision making behavior departs. Integration of these approaches within Group I may very well require a classification of aims if complementarity is to be achieved, or empirical test if the PDT approach continues its challenge to SEU theory and Edwards' BDT.

Group II approaches. None of these approaches are prescriptive. And although differences in intended function (to be discussed below) exist, it does not appear as if any one approach is prepared to deny the validity of any other approach with regard to its intended function (in contrast with Group I approaches). Thus, for example, although SJT declares its primary function to be that of description rather than explanation, it does not deny the validity of the explanatory efforts of IIT or AT as explanatory theories of judgment (methodological questions aside). Indeed, differences in intended function among Group II theories seem to be complementary. Thus, for example,

(a) AT intends to explain the direction and locus of causality (the distal variable of interest) in terms of the substantive aspects of the person-environment interaction. (One indication of this can be seen in Table II-D-1 which reproduces the first page of the subject index of "New Directions in Attribution Research," Harvey, Ickes, Kidd, 1976).

(b) IIT intends to find the rules, or the cognitive algebra, whereby information regarding persons and circumstances are

Table II-D-1

Subject Index

A

- Ability and knowledge, 413-416
- Ability/effort/chance attributions
 - and helping behavior, 320-331
 - and male-female performance, 339-341, 343-351
 - and reinforcement, 105-108, 113-114, 117
- Achievement motivation, 345
- Action sequences in behavior perception, 236-239
- Actor-observer differences, 41-43, 61-62, 88-89, 186-189, 294-297, 401-402
- Actor-partner differences, 362-365
- Affiliation, 209-210
- Aggression
 - and reciprocity, 250-254
 - and social influence, 304-307
- Altruism, *see* Helping behavior
- ANOVA cube, 406-417, 431
 - and correspondent inference theory, 411-447
- Anxiety, 148-154, 158-161, 405
- Arousal
 - and dissonance, 200-215
 - effect of, on task performance, 234-235
 - and emotional states, 143-164
 - of reactance, 249-250
- Attention, theory of, 179-182
- Attitudes
 - attribution of, and correspondent inferences, 394-398, 406

Attitudes (*contd.*)

- change of, 203-204, 206-214
 - and dissonance, 201-204
 - effect of feelings states on, 271-289
 - and perceived freedom, 75-76
 - and social influence, 292 ff.
- Attribution, Heider's definition of, 4

B

- Balance principle
 - and attitude inference, 271-287
 - and relation to attribution, 10, 13-14, 16-17
- Behavior modification, *see* Reinforcement
- Bias, attributional, 322-323, 428-429, 438
- Blame, *see* Responsibility
- Bogus feedback, *see* Feedback, false
- Break points in behavior sequences, 227-231, 238, *see also* Unitization of behavior

C

- Category-based expectancies, *see* Expectancies
- Causal schemata, 271-272, 427
- Causal taxonomies, 259, 312-316, 337, 358-362, 383-385, 406-408
- Causal unit formation, *see* Unit formation
- Chance outcomes
 - explanations for, 105

organized (or integrated) into an impression or judgment regarding a distal variable (see, for example, Anderson, 1974, p. 2), thus offering a formal mathematical description that provides the explanation of the cognitive process involved in person-environment interaction. And Anderson finds at least some verbal formulations of AT to be confirmed by empirical test.

(c) Since SJT intends to describe the behavioral consequences (e.g., conflict) that follow from the formal characteristics of judgment tasks or situations (without imputing a lawful character to them), its aims are not challenged by the other approaches.

In short, whatever differences in aim might exist among Group II theorists, they do not charge that any one of the other approaches cannot possibly fulfill its intended aim, nor that any other theory is grounded in logically false or empirically untrue premises. There even seems to be a certain recognition of complementarity of aim, if disparity in method, among Group II approaches. Yet it must be said that the present coexistence is only that, and that the relations among these approaches might best be described by "aloofness," with the major exception of Anderson's effort (and Mumpower's unpublished paper).

Integration of Group I and Group II Approaches

There is no mistaking the fact that there are differences between Group I and II with regard to intended function. The critical question is: Are these differences complementary, and, if so, does their complementarity serve to extend the domain

covered by judgment and decision research? Perhaps equally important: Do the differences require resolution?

The possibility of complementarity is raised by the persistent use of different terminology by Group I and II approaches. Even a cursory glance at the language of Group I and Group II theorists shows that there is a far greater use of the terms "decision," "choice," and "preference" in Group I and a much greater use of "judgment" and "inference" in Group II. Is this a difference that is significant? Do these terms imply Group I and Group II theorists intend to study different aspects of cognition?

Knowing and choosing. No conventional distinction between the aims of "decision theorists" and "judgment analysts" exists, so far as we know. Bock and Jones (1968) did draw what appears to be an appropriate distinction, thus:

Traditionally, the concept of a sensory continuum applies to the subject's experience of a continuously variable physical event which is called the stimulus. The subject's response in comparing two such stimuli is usually called a judgment rather than a choice. This implies that the subject has no personal preference towards the outcome of the trial, but is merely an objective observer. This terminology carries into applied work when, for example, we call the expert sensory tester a "judge."

This effort toward discrimination was obviously doomed because of its arbitrariness; why couldn't a response comparing two stimuli be called a choice, rather than a judgment? The need for the distinction apparently occurred to Slovic and Lichtenstein in 1971, but they rejected it as a "tenuous" one. In the second sentence of their well-known article they state that the "distinction between judgment and decision. . . will not be maintained here; we shall use

these terms interchangeably" (p. 16 in Rappoport & Summers). And in their subsequent review, Slovic, Fischhoff, and Lichtenstein (1977) did not address the distinction between the two terms at all. Zeleny, a decision theorist, went further than ignoring the distinction; he argued that "It is misleading (*italics ours*) to try to maintain the customary differentiation between human judgment and decision making." Although Zeleny (1977) does not indicate why he believes the distinction is "customary," he indicates that it

leads to unhealthy separation and inbreeding in both disciplines where an interaction would lead to mutual enrichment and rejuvenation. Decision making can be characterized as purposeful judgment, the judgment directed toward achieving a particular goal or goals. The difference between judgments and decisions is purely terminological--they both represent a process of making a choice (*italics ours*) among available alternatives. Whether the purpose of such a choice is an action, an expression or preference, or an optimization of some criterion, it will influence the way in which the corresponding judgment is formed. This is to say that judgment as a form of information processing is not being made per se, unpurposefully, with no specific objective or criterion in mind.

But if Zeleny was prepared to find the distinction between judgment and decision to be "misleading" and reduce them both to choice, he and Starr thought it to be important to distinguish between the terms "decision" and "choice" (Starr & Zeleny, 1977).

Thus: "when mostly attributes are involved, we tend to refer to such situations as those of a theory of choice, while the cases dealing with objectives may be referred to as a theory of decision making" (*italics in original*, p. 14). But these distinctions disappear because "In reality both the attributes and the objectives are often involved in a mixed fashion. We shall refer to both of

these categories as criteria. Criteria are both the attributes and the objectives judged to be salient in a given decision situation."

The title of R. N. Shephard's talk at the AAAS meetings in 1962 ("Use of judgments in making optional decisions") provides a further illustration of the uncertain status of the distinction between these terms: Could that title have just as well read: "Use of decisions in making optimal judgments"? And it is indeed curious to find Anderson (1974, p. 2) explaining that greater interaction has not occurred between IIT and AT because "much of the work on integration theory has been outside social psychology, in decision-making (*italics ours*)."

It seems clear that however convenient Slovic and Lichtenstein might have found it in 1971 to dismiss the distinction between these terms as "tenuous," the question of different referents for them needs to be examined more seriously than has been attempted so far. We shall not make that examination here, but we shall make the distinction. We do so because we believe that it will serve the cause of the development of a cumulative scientific discipline by indicating that the work of researchers on both sides of the conceptual watershed may very well be complementary. Specifically, we believe that integration would be better served by instituting that distinction because it will be useful in identifying the intended functions of the theories. For example. Theories of choice, decision and preference (Approaches 1, 2, and 3) aim at describing and explaining decisions to select, choose or to prefer one alternative (an object, an action, etc.) rather than another

AD-A061 750

COLORADO UNIV BOULDER CENTER FOR RESEARCH ON JUDGME--ETC F/G 5/10
THE COLORADO REPORT ON THE INTEGRATION OF APPROACHES TO JUDGMEN--ETC(U)
OCT 78 K R HAMMOND, G H MCCLELLAND N00014-77-C-0336
CRJP-213 NL

UNCLASSIFIED

2 OF 5
ADA
061750



as a function of one's interest or goal. To be as specific as possible, DT aims at discovering the "nondominated" alternative(s) in a set of alternatives. DT thus intends to identify and then eliminate dominated alternatives from consideration. Once identified, "dominated" alternatives are of little or no interest to decision analysts. The intended function of the theory, then, is to prescribe, and aid in, the choice of the nondominated object (or course of action, etc.) that follows from the expression of the decision maker's utilities (as well as probabilities). When the choice is made, the task is finished; DT is not curious about why persons should be less than rational.

Bayesian research, on the other hand, does have this curiosity, and it intends to show that DMs often select, choose or prefer the alternative they shouldn't prefer, if their probabilities and utilities are what the decision maker says they are. BDT offers one "explanation" for this, namely, that people fail to use the appropriate likelihood ratio, or fail to revise posterior odds as much as they should upon the receipt of new information. Although BDT is clearly different from DT in its effort to describe the degree and general form of departures from accurate probability estimates, it does not pursue the investigation of inductive knowing beyond this interest. And therefore it can hardly make strong claims to function as a general theory of knowing, as against its strong claims to function as a theory of decision or choice or preference.

PDT can make a stronger claim to being a general theory of knowing as well as a theory of choice, simply because of the wider scope of its theoretical, explanatory concepts. And because PDT studies a wider range of judgment problems than one observes in BDT, it might be argued that greater contact is thus made with an empirical environment. Nonetheless, it is true that both BDT and PDT study knowing only within the context of the biases in probability estimates of the correct state of affairs "out there." And although such estimates can be involved in knowing what's "out there," and therefore learning what such biases might be are bound to be of considerable value to students of quasi-rational knowing, they do not exhaust the cognitive activity involved in knowing. But of course such information would be of critical value in any SEU approach.

The above distinctions between BDT Bayesian research and PDT (Kahneman & Tversky) are no doubt somewhat tenuous, but when the watershed is crossed to the Group II approaches it is clear that we are now dealing with an approach that emphasizes knowing or learning about the environment in contrast to making choices about alternatives or preferences. For judgment theorists see problems of choice, action, and preferences for various outcomes of the world as subproblems of the more general problem of knowing. The study of knowing is thus propadeutic to the study of choice. Group II theorists intend to develop a theory of knowing from which a theory of decision is produced. This highly general aim can clearly be seen in all three approaches within Group II, none of which ascertain

subjective probabilities directly from their subjects as a general practice, and none of which are interested in utilities or rewards, although all are perfectly prepared to accept or request probability estimates or utility judgments in any special case. The primary goal of judgment analysis, in short, is to describe and to understand quasi-rational inductive knowing, rather than quasi-rational choice or preference. And as we shall see below, this difference in aim leads directly to a difference in the manner in which researchers evaluate the performance of the subjects they study.

To summarize, one strong reason for maintaining the distinction between the terms decision and judgment is that it raises the possibility that the different forms of research that are associated with these terms may complement one another, and may thus broaden the scope of the new discipline. Moreover, understanding the nature of the contribution of each approach may lead to the cumulative development of the discipline. Although we cannot consider that possibility in depth in this study, it certainly should be pursued with care in future work. And, as we turn to the matter of the evaluation of subjects' behavior, we shall see that differentiating the contributions of Group I and Group II approaches in this regard also serves the possibility of complementarity.

Evaluating the subjects' behavior. Differences in the intended function of Group I and Group II theories become apparent when we observe their attitudes toward the evaluation of behavior. In Group I, evaluation of the subjects' decision behavior is carried out with reference to the decision behavior of a mathematical model

of rationality; the model is used to evaluate "correctness," or subjective optimality. This is also true for BDT's MAUT; no empirically correct answer need be available.

Bayesian research and PDT research move one step closer to empirical, as against logical, evaluation of the subjects' performance by calculating the correct answer to a decision problem through the use of a statistical model (Bayesian in the first case, and both Bayesian and frequentist in the second). Because the results of evaluating the subjects' behavior have almost invariably indicated sub-optimal performance, the use of decision aids of various kinds has been recommended. Most recently, "debiasing" of incorrect probability estimates has been attempted, in the hope that if such biases can be removed in cognitive probability assessments, then more nearly correct probability estimates (i.e., answers that are more nearly compatible with those produced by statistical models with strong logical and empirical claims to truth) will be produced by decision makers. Thus, the culmination of the aim of ascertaining and evaluating the correctness of human performance is the provision of a remedy for the inadequacy of that performance.

In contrast, on the other side of the conceptual watershed, SJT and AT evaluate subjects' performance in judgment tasks in terms of the achievement of an empirical criterion. (IIT foregoes evaluation as a general practice, see Shanteau and Phelps, 1977). Thus, the intended aim of SJT is to study the acquisition and use of empirical knowledge (under conditions of environmental ambiguity, including

uncertainty), rather than to study the approximation of human probability assessment to optimal probability assessment, as in the case of BDT and PDT. SJT's intended function, then, is to study cognitive adaptation to an ambiguous (uncertain) environment. Thus, for example, SJT studies multiple-cue probability learning in which the evaluation of a subject's achievement of an empirically correct answer is the critical component of performance. To be sure, that performance is analyzed and often described in terms of various statistical concepts; nonetheless the correct answer with which the subject's answer is compared is a tangible, empirical one, not necessarily provided through calculation as in the case of Group I approaches.

This point can be seen more clearly in SJT's studies of interpersonal learning and interpersonal conflict; the evaluation of performance is in terms of empirical achievement of an outcome (a second subject's response, and, of course, that response is not under the control of the experimenter). Again, this does not mean that statistical analysis cannot be employed after the fact in order to analyze the components of the subject's achievement.

The same is true for AT; there is no reference here to optimal choices, preferences, or probability estimates; although the experiments do evaluate the performance of subjects. Although empirical correctness is not always emphasized, it can always be used as a reference point. The performance of subjects can thus be evaluated by showing that empirically incorrect attributions (or inferences) are made under certain conditions of social structure. It is the

relation between (social) structural conditions and attributions (inferences) that form the basic materials from which laws of attributional behavior are to be constructed.

In short, the employment of different methods for evaluating the subjects' performance follows from the distinction between choosing and knowing made above, and thus leads to further (tentative) support for the hypothesis that the Group I and II approaches may be complementary.

But what about IIT? Is it not also complementary in form to Group I approaches? Possibly, but it is more difficult to be clear about IIT in this context, for IIT is sharply different from the SJT and AT in relation to evaluation of its subjects' behavior.

IIT has indicated little, if any, interest in evaluating the fit of various models of optimal behavior to empirical behavior. As indicated above, IIT is wholly disinterested in the model or process that the subject departs from and is solely interested in the model that "fits," or accounts for, the subject's response. Anderson (1974, p. 66) makes this distinction clear when he states: "In the case of the Bayesian model, there is no doubt that it is wrong," and "although various attempts have been made to 'explain conservatism. . .it is doubtful that there is anything to be explained, since the effect exists only in reference to an admittedly incorrect model." In short, since the Bayesian model is "wrong," it should be abandoned as a contender in what Anderson believes should be an effort to find the correct description or explanation of human cognitive activity (however much interest.

rational models might have for decision theorists). It should be added that IIT gives no significant attention to the evaluation of subjects' performance in terms of the achievement of empirically correct judgments. In short, with the absence of an evaluative component regarding subjects' behavior in IIT, this approach appears to supplement, rather than complement, either of the Group I or Group II approaches described above.*

Therefore, we offer, as a hypothesis, the suggestion that IIT provides valuable information that supplements and thus adds to the cumulative character of the new discipline, whereas other approaches complement and thus tend to complete, or round out, the character of the new discipline.

To sum up, the differences in intended function observed above with respect to deciding and knowing are reinforced by the observations that Group I approaches generally, if not always, use

* The distinction may well be important and useful, therefore we quote from the Random House Dictionary College Edition (1968):

COMPLEMENT, SUPPLEMENT both mean to make an addition or additions to something. To COMPLEMENT is to provide something felt to be lacking or needed; it is often applied to putting together two things, each of which supplies what is lacking in the other, to make a complete whole: Two statements from different points of view may complement each other. To SUPPLEMENT is merely to add to; no definite lack or deficiency is implied nor is there an idea of a definite relationship between parts: Some additional remarks may supplement either statement or both (p. 275).

different sources for evaluating the performance of their subjects' cognitive activity from those used by Group II approaches. Not only do differences in sources of evaluation suggest that the analysis of choice behavior is the predominant aim of Group I approaches, while the analysis of inductive inference regarding the environment is the predominant aim of Group II approaches, they also suggest that Group I and II approaches complement one another. Awareness of complementarity of these approaches, together with the knowledge of the supplemental contributions of IIT, should increase the likelihood that the new discipline will become a cumulative one.

Differential function of the single-system case. As pointed out earlier (Scope), the single-system case is the research situation that typifies research on judgment and decision making. It is only this case that interests DT, and occupies the greatest interest of BDT, PDT, and IIT as a situation for producing knowledge about judgment and decision making.

However, SJT is not interested in the single-system case as a situation for studying and producing knowledge regarding judgment, despite all the emphasis on the single-system case that has been laid at the door of SJT (particularly by non-SJT researchers in naming it the "policy capturing" approach; for an example, see Slovic, Fischhoff, & Lichtenstein, 1977). Work within the single-system case by SJT researchers is derived, or extrapolated, from work in the other, more complex cases involving MCPL, interpersonal learning and interpersonal conflict (cases which the other approaches tend to ignore because of their absorbing interest in

the single-system case). That is, having learned what they believe they have learned from studies in these cases, SJT researchers are prepared to extrapolate from their research in these cases to the single-system case as a descriptive (not explanatory) effort, in the interests of providing a useful remedy for the suboptimal cognitive activity they believe they have observed in studies of MCPL, etc. But SJT (or probabilistic functionalism) has not indicated that work in this single-system case is intended to be productive of general explanations of judgment and decision making behavior. Consequently, SJT researchers do not offer competitive descriptions that are based on the single-system case, in contrast with DT, BDT, PDT and IIT (and possibly AT).

A good example of the extrapolation of findings from cases other than the single-system to the single-system case (and others) can be found in the remedy or cognitive aid offered by SJT for the less than perfect performance of subjects in cognitive tasks. The remedy is offered in the form of what is called "cognitive feedback" (as distinguished from "outcome feedback") by SJT theorists. The extrapolation is straightforward. If subjects fail to perform optimally or adequately in MCPL tasks, and if it is found (as it is with respect to normal and psychotic persons) that the failure is due to the inappropriate use of relative weights or function forms relative to those in the task, then SJT advocates that the subject should be provided with pictorial, easy-to-grasp information about them so that improvement in achieving the empirically correct answer will become possible.

In addition, this approach has been extrapolated to the single-system case. That is, a person's judgment policy is determined (i.e., described in terms of weights, function forms and cognitive control). Persons' policies are displayed for them (cognitive feedback) so that a person may change his/her policy under perfect control at the console. In short, it was the results of research in the double-system case that were applied to the single-system case--just the reverse of other approaches. (More will be said about cognitive feedback below in connection with the link between cognitive models and environmental models below; see "Uses").

IIT is deeply committed to offering descriptions (whether these are elevated to the level of explanations or not) of cognitive activity in the single-system case, and it is the derivatives from this case that are to be applied to others (say, group decision making), precisely the opposite manner of approach taken by SJT. In any case, IIT is wholly disinterested in remedies; neither suboptimality, subachievement, nor remedies for these are of interest to IIT researchers; finding the correct model to fit the integrative process is the primary problem of interest to these researchers.

We are uncertain about the AT approach. For although AT researchers do not appear to be interested in the single-system case at all, this conclusion may very well be incorrect, since proponents may provide counseling to clients that is based on results from certain experiments. Since the precise form of the

counseling, if it occurs, is unknown to us, we shall assume that it is irrelevant to our present purposes. We should only point out that if such counseling does occur, the direction of extrapolation of results would, in general, be similar to that taken by SJT, namely from the more complex cases to the single-system case.

There is one indication of change in AT that may well bear on its interest in the single-system case. A recent article by Jones and McGillis indicates that AT may have taken an entirely new turn that will link it more closely to Group I approaches. These authors (in Harvey, Ickes, & Kidd, 1976, p. 404) offer a version of AT which they label "correspondent inference theory," and describe it as follows:

Correspondent inference theory is essentially a rational baseline model. It does not summarize phenomenal experience; it presents a logical calculus in terms of which accurate inferences could be drawn by an alert perceiver weighing knowledge, ability, noncommon effects, and prior probability. But the role of the theory has been as much to identify attribution bias as to predict precisely the course of the social inference process. In a similar vein, Edwards (1968) has stated regarding decision theories that "they specify what an ideal decision-maker would do and thus invite comparison between performances of ideal and real decision-makers [p. 34]." The theory cannot be invalidated by experimental results any more than game theory can be invalidated by the choices of players in a prisoner's dilemma game.

Note the references to Edwards; this statement indicates more than a recognition of common interests. Jones and McGillis go much further and suggest that they are producing a normative or prescriptive theory similar to decision theory or game theory. They also say that "correspondent inference theory has a different goal than attribution theory" (p. 406).

To summarize: the single-system case may be the research situation in which all approaches (except AT?) share a keen interest, but it is clear that this situation is studied very differently by different approaches because their aims are different. Equally important, the remedies that different approaches offer are different because they are directed toward different (putative) cognitive deficiencies.

Does this situation imply complementarity and broadening of scope? Or does it imply antinomies that must be resolved theoretically and/or empirically? We are not sure. Consider again Slovic's comments regarding PDT:

Subjective judgments of probability and utility are essential inputs to decision analysis. We still do not know the best ways to elicit these judgments. Now that we understand many of the biases to which judgments are susceptible, we need to develop debiasing techniques to minimize their destructive effects.

The complexity of that task is also indicated, however, thus:

Simply warning a judge about a bias may prove ineffective. Like perceptual illusions, many biases do not disappear upon being identified (p. 238).

Would it not be reasonable for proponents of Group II approaches to ask: If we still do not know the best ways to evoke subjective probabilities, and if they are subject to so many crippling biases when they are evoked, and if attempts at debiasing are likely to be difficult to achieve at best, and ineffective at worst, then why should not the efforts to aid the decision maker that are based on ascertaining subjective probabilities be abandoned in favor of other approaches? Perhaps subjective probabilities

(and utilities) are not essential inputs to the decision making process. Perhaps the study of knowing is propadeutic to the study of choice. Perhaps if one gains an understanding of knowing, in the manner of Group II approaches, the remedy for inadequate cognitive efforts will become, and indeed perhaps already has become, apparent.

E. Principal Concepts

Introduction

One of the primary reasons for attempting to integrate various approaches to human judgment and decision behavior is that many different concepts seem to be employed by different researchers to refer to the same phenomena. Moreover, the wide range in terminology gives rise to the uncomfortable suspicion that many authors may be using different words to refer to the same concept. Equally uncomfortable is the suspicion that the same word may be carrying different conceptual and operational meanings for different researchers. Does the Bayesian researcher's "datum" have the same meaning as the IIT researcher's "stimulus"? And does "stimulus" have the same meaning as the SJT's researcher's "cue"? Most important, does the concept of uncertainty convey the same meaning to all researchers in the field of "judgment and decision under conditions of uncertainty"? As we have seen earlier there is doubt as to whether judgment and decision should be used interchangeably. What about "uncertainty"?

In this section we will not try to settle such questions by references to operational definitions; these are discussed in the Procedure section. Rather, we shall consider at a conceptual level the principal concepts used by the six approaches.

Decision Theory

The major concepts in DT are, of course, decision, choice, preference, and probability, utility, and aggregation. Both probability and utility are treated as judgments that are produced in response to special circumstances (e.g., lotteries) presented to the decision maker by the decision analyst. Such circumstances are intended to permit the measurement of each in ordinal (if not interval or cardinal) form. Ideally, if a decision maker's preference is larger for one alternative than another, it will be as a result of the mathematical, analytical aggregation by the decision analyst of the separate probability and utility judgments made by the decision maker.

Consider first the probability judgment. Note that it is a dependent variable that is ascertained directly in response to specific circumstances. It is essential to note the role of probability (and thus uncertainty) as a dependent variable of primary importance in DT for two reasons: (a) it has been carried into the work of BDT and PDT in this form, and (b) it has not appeared in this form (except adventitiously) in Group II approaches. As a result, the role of probability as a psychological concept is different in these two general approaches.

Much the same can be said for the concept of utility. It is also a metatheoretical proposition within DT that the utility of an alternative exists in the mind of the decision maker and it is assumed that it is essential to obtain a direct measure of it. In contrast to the concept of probability, DT and all other

approaches employ the concept of utility in much the same way, although other approaches may not employ it in every situation.

The aggregation of probabilities and utilities is the cognitive activity of the decision maker that combines, integrates or organizes the probabilities and utilities associated with various alternatives. DT is prepared to use either additive or multiplicative methods of combination, depending upon certain conditions.

DT researchers have indicated considerable interest in scaling and transitivity, and indeed, their interest in these matters attracted the attention of Edwards and other psychologists (e.g., Coombs; see also IIT). In addition, the mathematical consequences of nonindependence interest DT researchers. (Keeney & Raiffa, 1976, devote some space to these problems [17 page references in the index, plus several other references to "conditionality"].) Nonindependence, rather than scaling, is apt to catch the attention of psychologists interested in the question of how people cope with "causal texture" (cf. SJT and AT).

The importance of the concept of independence-dependence can also be seen by consulting the Index to Keeney and Raiffa (1976) and noting the entries under "Assumptions": These are

additive independence, 230, 295

conditional independence, 333

group preference, 536

parametric dependence, 259

preferential independence, 101, 284

state dependence, 214, 502

utility independence, 226, 284

verification of, 191, 276, 382, 398, 452

additive independence, 263

preferential independence, 299

risk aversion, 192

utility independence, 264, 299

Perhaps the most important development in DT is with regard to decisions involving multiple, rather than single, objectives. Objectives, in turn, are defined in terms of their attributes. Thus, "the Postal Service objective 'minimize total transit time for a given category of mail'. . . was measured in terms of the attribute 'days'" (Keeney & Raiffa, 1976, p. 34). As a result multiple attributes must be considered, and, therefore, nonindependence; in short, trade-offs must be expected. All of these concepts are also found in BDT.

Behavioral Decision Theory

DT and BDT share essentially the same mathematical logic and the same concepts (decision, choice, preference) as major end-terms, as well as sharing the concepts of subjective probability and utility (as dependent variables) and, of course, both include aggregation as a concept that refers to the combination of them. And as in DT, BDT takes it as a metatheoretical proposition that uncertainty exists in the mind of the decision maker and, therefore, it is essential to obtain a direct measure of it. This is done by requesting Ss to express their uncertainty in terms of a probability scale, odds, a likelihood ratio etc. Explanations

for, or descriptions of, the source(s) of uncertainty in the mind of the decision maker are not an integral part of the BDT approach and, therefore, no detail is provided by BDT on this matter (although there is some indication that if such explanations were offered they would be similar to those offered by SJT; see Edwards' introduction to Slovic & Lichtenstein, 1971).

BDT also accepts DT's metatheoretical proposition regarding the need for the concept of utility and attempts to measure it (as Edwards put it: "what's at stake and what are the odds"?), but without the detailed concern traditionally exhibited by DT.

All, or almost all, of the concepts mentioned above in connection with DT have been investigated empirically by BDT. Note that these concepts are inherited from economic theory and Bayesian probability theory, not psychology, and, therefore, empirical work should carry implications for the use of the theory by decision analysts--or so the BDT's think. Thus, in empirical Bayesian research, we find such concepts as conservatism, the diagnostic impact (of a datum or event), conditional dependence, unconditional hypothesis, independence of data, (mis)perception, (mis)aggregation, and diagnosticity, etc. All these terms were developed in the course of the research on departures from Bayesian optimality; e.g., why the conservative bias? In order to see the descriptive function of these terms we briefly consider the typical Bayesian research paradigm.

The subject is required to observe a datum or event (drawn from a population of data or events), make an estimate of the

probability that it was produced by a given source; observe a second datum, at which point the subject revises whatever prior probabilities (or odds) he held prior to observing the datum, the extent of that revision providing a measure of the impact of the datum on his posterior probabilities (and thus the subject's evaluation of the diagnosticity of the data), in addition to which the subject may aggregate all the information he has acquired up to this point and instead of revising his posterior probabilities may be requested to evaluate the likelihood ratio regarding the population which produced the event. Semipsychological concepts have been introduced into those Bayesian experiments where learning is involved (e.g., pay-offs and feedback) both of which were introduced in the effort to discover whether subjects could reduce their "conservatism" and thus behave (i.e., revise their estimates) more nearly in accord with the "behavior" of the (optimal) Bayesian model. None of these concepts are in jeopardy today in the sense of being considered wrong or useless, as for example "fractional anticipatory goal response" might be considered. But their value as descriptors of significant psychological processes is being questioned by PDT and IIT.

[MORE]

Psychological Decision Theory

This approach also includes the concepts decision, choice, and preference, as well as subjective probabilities and utilities because the point of departure for this approach is also SEU theory.

In addition, we find the term "heuristics," used to describe various rules of thumb used by subjects to decide the correct answer to a problem involving a statistical inference. And it may well be that it is these heuristics that interfere with the rational use of information, and for which DT seeks to provide a remedy.

Uncertainty continues to be conceived of as a pervasive state in the mind of the decision maker, and PDT also assumes that this state of mind also is to be ascertained directly in the form of a probability estimate (or choice) in relation to a set of task conditions, although, as we shall see, these conditions are far different from those employed by DT or BDT. Similarly, no detailed theoretical analysis is offered as to why such uncertainty should exist. Judgment tasks are described in statistical terms, although no special group of statistical concepts is emphasized (as in the case of BDT's predilection for Bayesian concepts); rather, PDT theorists rely on the entire repertoire of statistical theory and employ various methods to evaluate the optimality or correctness of any given heuristic putatively used by a decision maker.

The major significance of PDT (and the reason why we have assigned it the name we have) is that in this approach we meet for the first time the use of psychological (in contrast to statistical) concepts to explain the departure of the behavior of human decision makers from optimality. These concepts, moreover, are not transported from research conducted in other areas and applied post hoc (e.g., "cognitive strain") but are derived directly from research

on judgment and decision making. Included are such concepts as availability, representativeness, and anchoring. Some of these terms have been used before in traditional psychophysics (anchoring), in cognitive psychology (availability; cf. Bruner, Goodnow & Austin, 1956), and some are new (representativeness).

One of the more recently introduced explanatory concepts is that of causal schemata, thus, "the impact of evidence (cf. "impact of a datum" in Bayesian research; au) on intuitive judgments of probabilities depends critically on whether it is perceived as causal, diagnostic or incidental" (Tversky & Kahneman, 1977). The use of statistical referents (see II/C above on evaluation) is illustrated by the subsequent statement: "base-rate information (*italics ours*) which is given a causal interpretation affects judgments, while base-rate information which cannot be interpreted in this manner is given little or no weight" (*italics ours*). Thus the above statement not only signifies PDT's employment of statistical logic (i.e., base rate) to evaluate a subject's use of information, it also indicates the universal ($n = 6$) resort to the psychological concept of weight; apparently no approach can do without it, although AT makes uncertain use of it. (Equally significant, no approach has given attention to its theoretical status, except possibly Kahneman & Tversky in Prospect Theory.) The employment of the concept of weight by PDT thus provides a link between PDT and other approaches. For although the independent variables (or factors) that are believed to affect judgments of probability are different from those considered by DT and BDT,

nevertheless, the critical question in PDT is still directed toward the differential "impact of evidence on intuitive judgment of probabilities." In other words, how much weight for a given "input" of what type? (This point is developed further in the Integration section.)

As noted above, the independent variables that PDT considers to be determinants of "differential impact" or evidence, or to determine the "weight" that is assigned to different varieties of it, constitute a different class of variables than have been used by DT or BDT. The major difference is that the data presented to the DM is described by PDT not only in formal (i.e., statistical) terms, but psychological ones. For example, "the ordering of events by their subjective probabilities coincides with their representativeness" (p. 431, Kahneman & Tversky, 1972). Representativeness (a psychological concept) therefore explains the magnitude of subjective probability judgments. But what is "representativeness"? This term is defined in terms of the degree to which an "uncertain event or a sample is i) similar in essential properties to its parent population; and ii) reflects the salient features of the process by which it is generated" (p. 431, Kahneman & Tversky, 1972) (*italics ours*). "Similarity" is given detailed treatment by Tversky (1977) in a paper devoted only to that topic but that material need not be discussed here. Clearly, however, "similar" and "salient" are subject-oriented terms.

Because terms such as "similar" or "salient" are subject-defined, they are members of a class of organism-centered

definitions. These terms are used to represent some of the variables that affect judgments of probabilities. And it is these variables that give a priori meaning to the concept of weight, and thus explain why this piece of evidence is ignored and that piece of evidence is not. In addition, different organism-centered terms introduced by PDT (seem to) belong to different classes. For example, representativeness appears to be an organism-centered definition of an object-attribute, whereas availability seems to refer to an organismic process (as does anchoring). (This point is also discussed in more detail in the Integration section.)

Transition

The shift in orientation between Group I approaches and Group II approaches may be seen in the changes in the names of principal concepts we shall observe in use as we move across the conceptual watershed. No longer are such terms as probabilities and utilities employed as primary theoretical concepts, and no longer is it taken for granted that these concepts should be translated directly into measurable quantities. Nor is it found necessary to attack the validity of the DT approach by showing that persons do not cope with subjective uncertainty as they should. In only one of the three approaches (SJT) is uncertainty treated as a major concept and it is given far different treatment here than in Group I approaches.

New terms will be encountered among Group II approaches; for example, cue, stimulus, percept, and information are used to

describe inputs to the organism, and concepts such as organizing principle, integration, inference and attribution are used to describe the cognitive activity of a "subject" or "organism" or "knower" rather than a "decision maker."

The concept of a watershed, rather than a gap or barrier, is particularly useful with regard to the principal concepts used by different approaches since the change from Group I to Group II is far from abrupt. Some terms (cue or stimulus, for example) are apt to have a broad range across nearly all approaches whereas others are highly restricted (e.g., ecological validity). The transition from Group I to Group II implies a continuous, if uneven, change rather than a sharp change in approach; the change is there, however, and it's recognizable.

Social Judgment Theory

SJT is the first approach to be encountered in moving across the conceptual watershed between Group I and II approaches because of the great emphasis on the concept of uncertainty that it shares with Group I approaches. As noted above, however, there is a clear difference in the function of this concept in Group I approaches and in SJT. For SJT postulates the emergence of uncertain (probabilistic) behavior as a function of an uncertain and generally ambiguous environment. And although SJT might find it advantageous in some circumstances to request subjects to make judgments about probabilities (paralleling the work of Group I approaches) such judgments simply constitute one instance of the

general class of judgments. (The same is true for IIT and would be true for AT if it requested probability judgments from its subjects.)

The concepts of value and utility (as these terms are used in Group I approaches) play almost no role in SJT. As in the case of probability judgments, utility judgments are simply one member of a class of judgments. Far more emphasis is placed on achievement, that is, the correct attainment of the distal variable, and that emphasis points to the fact that the fundamental conceptual paradigm for SJT is not the single system case (as it is for DT, PDT, and IIT) depicted in Figure II-C-1, but the double-system case in Figure II-C-2.

Uncertainty, therefore is inferred from the semierratic behavior of the organism as it attempts to cope with a semierratic task; it is not ascertained directly by requiring the person to make a judgment of probability values, or choices between lotteries (see the large series of studies carried out by Brehmer and his colleagues on MCPL). If probability judgments are requested at all by SJT (or IIT, or AT) such judgments are treated merely as another instance of a general class of inferences about a not-directly-observable state; uncertainty is inferred from inferential behavior in this case as well as others. In short, when Group I investigators study uncertainty, they use techniques to evoke persons' subjective feelings about their uncertainty, whereas when Group II investigators study uncertainty they observe its occurrence in the behavior of their subjects. So far as we know, the

significance of this distinction in the role of probability and uncertainty in judgment has never been explored.

The following paragraphs taken from the most recent statement of the general principles of S-JT are provided for the sake of convenience.

BASIC CONCEPTS

RELATIONSHIPS: THE FUNDAMENTAL UNITS OF COGNITION

The fundamental concept ordinarily employed to describe an environmental "input" to the organism is the stimulus. That concept is rejected here. Although both Tolman and Brunswik used this term, they did not make a complete conceptual commitment to it; both argued that the objects and events apprehended by an organism do more--and less--than "impinge" upon it. Not only does the organism cognitively act on the "input," but the perceived object carries implications for *other objects*. That is why Tolman's position was labeled an S-S theory (that is, a "sign-significate" theory) and contrasted to an S-R (stimulus-response) theory by competing theoreticians of his time. And that is why Brunswik used the word "cue" to refer to various dimensions of the perceived world. Both these terms, "sign-significate" (or as Tolman also put it, "sign-Gestalt") and "cue," have in common the notion that the raw materials of perception point outward from the organism toward various aspects of the person's ecological surroundings. And whereas "sign-significate" and "cue" point *outward* from the organism to the environment, the concept of stimulus points *inward*. It is for this reason that S-R theories in general do not include concepts relating to the environment and that S-R judgment theories, in particular, do not include concepts referring to the properties of judgment tasks (see, for example, Anderson, 1971).

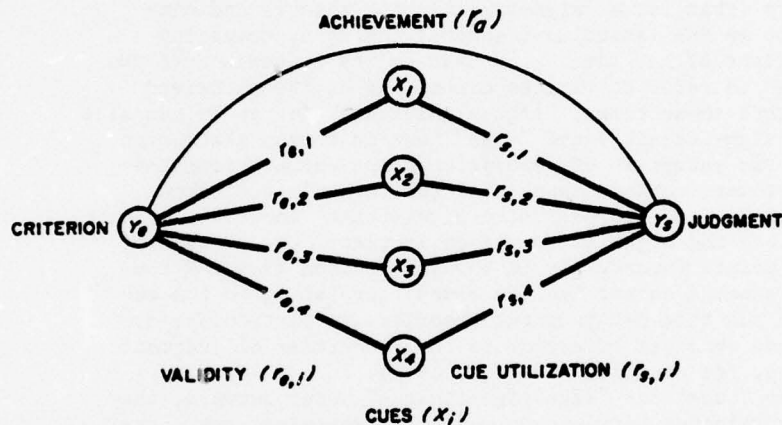
Because "cues" and "sign-significates" point outward, they involve a relation between two variables--proximal and distal, the given and the inferred. Choice of that relation as the fundamental unit of cognition has profound consequences, of course, and it was this choice that eventually led Tolman to introduce the concept of the "cognitive map" in 1948; he argued that cognition involves a subjective representation of the interrelations of goal paths in the organism's environment. Brunswik went further; he demanded a more detailed analysis of the *environment* and a less detailed analysis of the *organism*. Thus, for example, he remarked:

Both organism and environment will have to be seen as systems, each with properties of its own. . . . Each has surface and depth, or overt and covert regions. . . . It follows that, much as psychology must be concerned with the texture of the organism or of its nervous properties and investigate them in depth, it must also be concerned with the texture of the environment [1957, p. 5].

Brunswik's admonition to psychologists to "be concerned with the texture of the environment" gives clear direction to the student of human judgment; his first step must be to learn about and to understand the texture (and by that we mean the causal ambiguity) of the relationships among variables in the tasks which require human judgment. (The methodological corollary is that such ambiguity among relations must be represented in the judgment tasks used to study human judgment.)

PRINCIPLE OF PARALLEL CONCEPTS

As can be seen in the above quotation, Brunswik indicated that organismic and environmental systems should be described in symmetrical terms. That symmetry is represented in what Brunswik called the "lens model" of behavior indicated in Figure 1. (Space does not permit more than a cursory reference to the conceptual implications of the lens model; the best of several original sources is Brunswik's "The Conceptual Framework of Psychology," 1952; a secondary source which presents part of what is contained in several original articles is Hammond's *The Psychology of Egon Brunswik*, 1966.)



Brunswik's lens model.

As Brunswik describes the lens model, it becomes clear that he employs a principle of parallel concepts, for each concept on one side is paralleled by a similar concept on the other. Thus, cues on the task, or ecological, side vary in *ecological validity*, and on the organismic side there is variation in *cue utilization* by the subject. And just as the relations between cues and distal variables on the ecological side may assume various (linear, curvilinear) forms, according to the principle of parallel concepts, the relations between cues and judgments may also assume various function forms on the organismic side. The investigator has similar interests with regard to both sets of variables: to what extent ecological validities are matched by cue utilization and to what extent ecological function forms are matched by subjective function forms. Social judgment theorists are also concerned with the extent to which the

principles of organization that control the task system are reflected in the principles of organization that control the cognitive system of the subject.

It is the principle of parallel concepts, therefore, that produces the symmetrical relation between the descriptive terms applied to the organismic system and to the environmental system, and it is this principle that is responsible for the fact that Social Judgment Theory (SJT) includes a set of concepts which apply to task systems as well as person systems.

DISTINCTION BETWEEN SURFACE AND DEPTH

This distinction is essential to SJT. It derives from the proximal-distal separation in perception theory and thus refers to the separation between what is given and what is inferred. *Surface* data are (given) cues to (inferred) *depth* conditions in the judgment task. By virtue of the principle of parallel concepts, this distinction also applies to organismic judgment systems (see Figure 1). Separation of surface and depth is critical to any theory of judgment (or inference), for it raises the question of the properties of the region that intervenes between them. Because of the importance of this region, we have named it the *zone of ambiguity*.

The Zone of Ambiguity

The region between depth and surface variables in a given judgment task involves the relations between cause (depth) and effect (surface). Because a single effect may be produced by several causes, as well as because multiple effects may be produced by a single cause, there is ambiguity from cause to effect and effect to cause. Because causes may be related, and because effects are interrelated the network of task relations can be said to be entangled. Moreover, causal ambiguity is produced because (1) surface data are less than perfectly related to depth variables, (2) functional relations between surface and depth variables may assume a variety of forms (linear, curvilinear), and (3) the relations between surface and depth may be organized (or combined) according to a variety of principles (for example, additivity or pattern). These circumstances give more specific meaning to the term "causal texture," or causal ambiguity.

In short, causal ambiguity within the zone of ambiguity is the source of the human judgment problem, as well as a source of the misunderstandings and disputes that occur when judgments differ.

Information Integration Theory

This approach takes us further in the direction of a traditional psychological approach. Measurement of probabilities, utilities and the analytical aggregation carried out by the decision analyst are activities far removed from the principal business of IIT, that of discovering lawful relations between stimulus and response in circumstances involving cognitive activity. The character of IIT can be seen in the names of its principal psychological concepts. These are stimulus, response, information, integration, valuation, and weight. The concept of discounting takes on a special significance because it relates to (conflicting) cue interdependency. Thus (as in the case of SJT), Information Integration Theory omits subjective probabilities or utilities from its fundamental descriptive terms, nor does it argue for the superior logical status of any specific aggregational rule, organizational principle, or integrative mechanism to be applied to the weighted, subjective scale values of stimuli. Information integration mechanisms are given a prominent place in this approach and the generic term cognitive algebra (e.g., adding, averaging, multiplication, etc.), is used in contrast to some single analytical

mechanism such as the Bayesian rule in DT, or the pervasive additive principle in MAUT. Additionally, IIT refers to responses of the organism or subject (not a decision maker) which must be subject to the same scaling analysis as are stimuli.

IIT may be further discriminated from SJT by IIT's omission of uncertainty (and its derivative, probability) from its list of fundamental concepts. Thus IIT moves further from General Approach I than SJT in that (a) uncertainty is not a fundamental part of its conceptual system, (b) it does not emphasize inductive inference as an epistemological function, nor (c) include detailed consideration of the environment as a source of cognitive uncertainty.* When IITs use the term judgment it is in its psychological sense (as in judging the weights of hefted objects) but judgment is apparently also intended to carry the same meaning as "inference," as in the inference of a distal, impalpable state of affairs. Additionally, the term "impression" is often used to refer to the more general (distal) cognitive activity of inference (thus, "impression formation" in studies of forming impressions of personalities), but IIT apparently sees no compelling reason for distinguishing among the terms judgment, impression, inference or attribution; and the terms choice, or preference are simply considered to be instances of inference.

* Unhappily, life is multidimensional for judgment and decision researchers also as well as for their subjects; there are other dimensions on which IIT is closer to Group I approaches than SJT.

[MORE]

Attribution Theory

The key concept, of course, is attribution. As indicated above this term is directly translatable into a special case of the term inference or judgment (see Harvey, Ickes, & Kidd, 1976, p. 4 for Heider's definition of attribution). It is important to note that choices of action or preference for friends, say, follow from different attributions; AT is, however, primarily concerned with causal attributions (or, inferences about causality).

The concept of probability does not play a large role in this approach. Indeed, although the originators of AT (Heider) and SJT (Brunswik) shared not only a mutual intellectual respect (and friendship), frequently cited one another's work, and were particularly interested in "causal texture," probability is simply subjective uncertainty arising from ambiguity in Heider's theory. It does not receive theoretical treatment as it does in Brunswik's work. "Probability" for example is not indexed in Heider's book, nor is "uncertainty," or "weight," although "mediation, ambiguous" is. But it is organizational, not statistical structure that creates ambiguity for Heider, not relative frequencies of the co-variation or co-occurrence of events (see pp. 35-44 in Heider, 1958). And the same is true for contemporary AT. Although one might say that one finds in Heider a "grudging" acknowledgement of the "grain of truth" in Brunswik's probabilistic argument, that acknowledgement is overwhelmed by the return to the search for

invariances in the environment (see especially Heider, p. 44); the search for psychologically defined cues that are "structurally safe" (our term) and thus afford stable reference points in social circumstances. And despite numerous opportunities for contemporary attribution theorists to link their arguments to judgment and decision research, such links almost never occur, mainly, we believe, because of the probability approach taken by these researchers, a conclusion also reached by Fischhoff, 1976. (For a particularly clear example of a missed opportunity, see Wortman in Harvey, Ickes, & Kidd, 1976; for examples of links, see Jones & Davis, 1965; Jones & McGillis, 1976; Fishbein & Ajzen, 1975.) Utility is a different matter; it is given considerable importance by AT, but it is discussed in terms of valence or value.

Despite his aversion to a probabilistic approach, Heider is keenly aware of the role of interdependency among cues in creating ambiguity. In a section entitled "Economy of interpretation--redundancy," he says (pp. 51, 52):

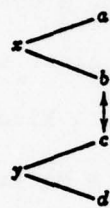
Economy of interpretation—redundancy. We have seen that a stimulus which is ambiguous as long as it is given singly, may become unequivocal with the addition of further data. It is important to stress that this specificity is established through the meaningfulness of the integrated perceptual field. But of two equally meaningful integrations, the one that is less complex, the one that requires fewer assumptions, fewer data in general, seems in general to be preferred. This is sometimes referred to as the principle of parsimony, a principle well known in the philosophy of science, and which may have its analogue in perception.

It is sometimes said that the objective of science is to describe nature economically. We have reason to believe, however, that some such process of parsimonious description has its beginnings on a fairly naive perceptual level. . . . It appears likely that a major function of the perceptual machinery is to strip away some of the redundancy of stimulation, to describe or encode incoming information in a form more economical than that in which it impinges on the receptors. (Attneave, 1954, p. 189.)

Now to illustrate economy in perception. Let us assume that stimulus x is ambiguous, that it could be interpreted as mediating either a or b , two different features belonging to the distal sphere, i.e., the environmental world. In the same way stimulus y could be seen as b or c . If x and y are given together, they can be "explained" by the hypothesis " b is there," or by the hypothesis " a and c are there." The first hypothesis is "cheaper"; it refers the stimuli to only one underlying entity, whereas the second hypothesis assumes two entities



The same principle can be applied if two of several meanings underlying ambiguous data imply each other. For instance, suppose that x means a or b , that y means c or d , and that b and c imply each other mutually. Then, if x is interpreted as b , it also transmits the information that c is there, since b implies c . Or, it is a simpler hypothesis to interpret xy as bc than as ad , since a and d are two independent facts and bc makes an integrated group and contributes only one fact according to the restrictions of the system. It will be seen that the first example above in which two meanings are identical is really just a special case of this one.



(Heider's discussion of one hypothesis being "cheaper" than another and thus more likely to be used can be related to Tversky and Kahneman's "Causal Schemata." Also recall the extensive reference to "independence" in Keeney & Raiffa's assumptions mentioned above.)

Curiously, the concept of cognitive organization (or aggregation or integration) is not indexed by Heider, nor does it seem to be singled out for analysis by AT, although one somehow gains the impression that the process of cognitive organization has a large role in AT, possibly because of the emphasis on "structure."

In the second phase of the development of AT we find that Kelley gives considerable attention to both interdependency (as we noted was the case with Heider) as well as the integrative process (not addressed as a major topic by Heider [but cf. Anderson's assistance provided for AT in Berkowitz, 1974]).

Definite recognition of the entangled nature of task dimensions can be found in Kelley's (1973) statement on attribution theory. For example, "our attributor. . .sometimes treats. . . causes. . .as being interdependent. . ." as do scientists, although the latter "know that causes do not occur independently and in all combinations. We conduct experiments precisely for the purpose of creating such circumstances. . . .At the same time, we know that such separation and independence are not characteristic of real-life. . . .It seems reasonable to assume that the lay attributor shares this awareness of the possible interdependence among causal factors. . . ." Moreover, "if and when the layman makes such assumptions, the inferences he will draw. . . will shift in a drastic manner" (italics ours).

In Kelley's version of the theory, persons cope with the difficulties of entangled task dimensions by applying the principles of the orthogonal components design (called the "ANOVA cube" by Jones & McGillis, 1976). Thus, Kelley asserts that the orthogonal components design is the basis of ordinary human inference, or "common sense." That is, (unaided) human inference

follows, or attempts to follow, the logic of modern experimental design (in effect, Mill's Canons) in coping with the inference problems of everyday life.

There is some recognition by Kelley (p. 31 ff.) of the asymmetry between the nature of the inference task and the cognitive processes he believes are brought to bear on it. We are left, however, only with the warning that the asymmetry "poses a problem for a complete theory of the attribution processes" and its associated methodology.

Since Jones has been a leading contributor to AT, it should be mentioned that he and McGillis present a Chapter in Harvey, Ickes, and Kidd (1976) entitled "Correspondent Inferences and the Attribution Cube." They indicate that "correspondent inference theory" is "essentially a rational baseline model" (p. 404 in Harvey, Ickes, & Kidd) and thus bring it near to Group I approaches. The Figure II-E-1 on the page following indicates that Jones and McGillis are prepared to see the attribution process in probabilistic terms closely associated with the Bayesian view. If "correspondent inference theory" as depicted in Figure II-E-1 were to become an integral part of AT the latter would be changed significantly.

Integration

Integration within Group I approaches. Integration within these approaches will require some direct efforts by those working within the framework of decision analysis to come to grips with what appear to be significant antinomies. On the one hand, Keeney

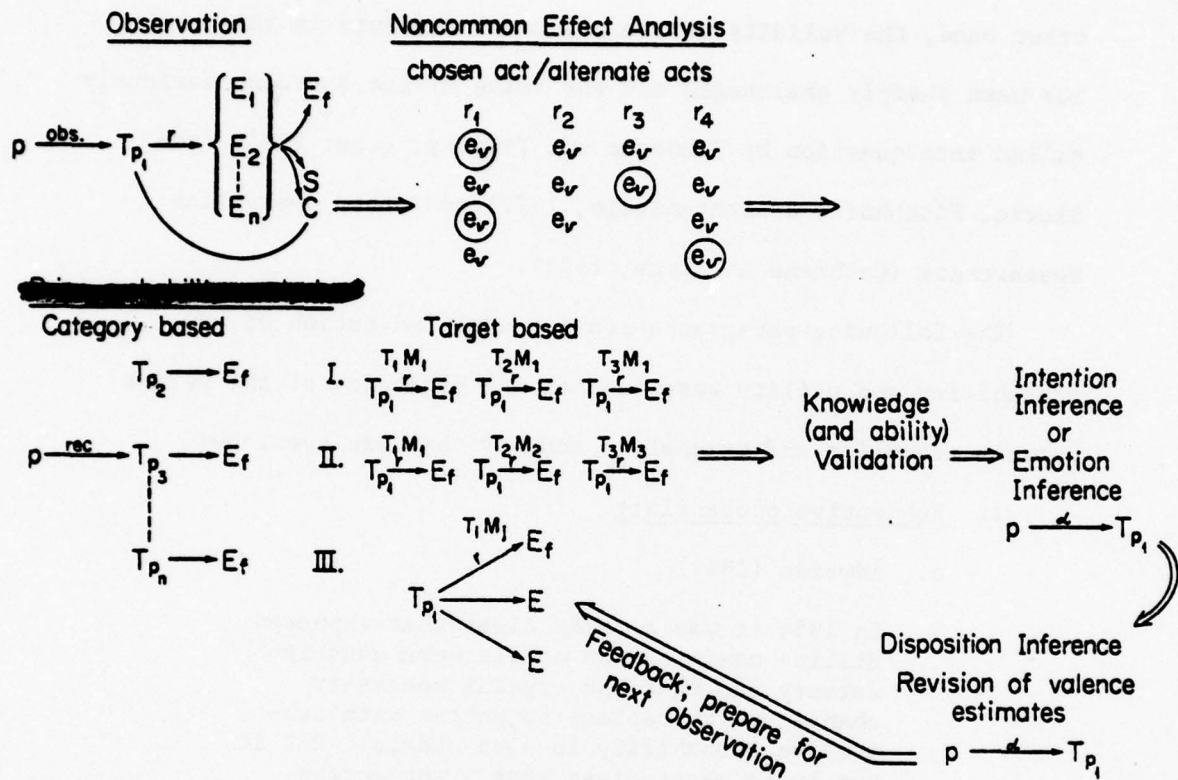


FIG. 3. Action sequence in an integrated attributional framework. Key: p , perceiver; T_p , target person; e , effect; \odot , noncommon effect; obs., observe; rec., recall; α , attribution; T , time; M , modality; E , entity; E_f , figural entity; S , setting; c , circumstance; v , valence.

and Raiffa's 1976 book was awarded a major prize by the Operations Research Society of America (ORSA), and thus endorsed as a highly valued representation of decision theory by their peers. On the other hand, the validity of the principal concepts in this book has been sharply challenged and the value of the approach seriously called into question by Kahneman and Tversky, other PDT's (see Slovic, Fischhoff, & Lichtenstein, 1977) and other operations researchers (Cochrane & Zeleny, 1973).

The following paragraphs provide some indication of how probability and utility were seen at the beginning of the merger between economics and psychology and how they are seen now:

1. Subjective probability

a. Edwards (1961)

In 1954 it was already clear that expected utility maximization models were unsatisfactory and that the crucial necessary change was to replace objective with subjective probability in such models. But it was by no means clear what a subjective probability is. In 1960 it is clear what a subjective probability measure is, but it seems unlikely that subjective probabilities conceived as measures are any more adequate than objective probabilities in the face of the data. Less restrictive definitions of subjective probability, which do not require them to be measures in the sense of measure theory but which still preserve a form of the SEU model, are in much the same state of ambiguity and ill-definedness as in 1954.

b. Slovic (1976; p. 238)

Subjective judgments of probability and utility are essential inputs to decision analyses. We still do not know the best ways to elicit these judgments. Now that we understand many of the biases to which

these judgments are susceptible, we need to develop debiasing techniques to minimize their destructive effects.

2. Utility

a. Edwards (1961; p. 67)

In 1954 there was general agreement on what utility is and that it can be measured, but no real agreement on how to measure it. In 1960 there is at least a conceptually adequate method (assuming a SEU model)--but some doubt about whether a SEU model, and therefore any measurement methods based on it, can stand up to these facts.

There is no indication by Raiffa in 1968 that the concept of utility as he employs it (as a choice between lotteries) "can stand up to these facts." Nor is there any concern about this indicated by Keeney and Raiffa roughly 10 years later. Should there be?

b. In 1973, at least some decision theorists thought decision theorists ought to pay attention to the "facts" produced by psychological research. Thus, Cochrane and Zeleny in their introduction to the proceedings of the 1972 conference wrote (p. xiii):

The results of psychologists and other social scientists--that individuals and groups do have multiple goals, none of which can be pursued to the complete detriment or sacrifice of the others--have been ignored. Even further, these goals often change not only in the course of time but also in the course of their pursuit. Finally, these goals are certainly not independent of the means used to pursue them. The traditional utility approach, suggesting that the alternative with the highest utility will be chosen, is

not of much help under such conditions. About the only property this simple idea has is transitivity. Transitivity, however, is not the behavior people manifest when facing situations characterized by multi-attribute consequences.

The references to "the results of psychologists and other scientists" and the reference to "the behavior people manifest" suggest that these authors believe that there is an interface between DT and the research carried out by behavioral scientists.

But it does not appear that the issues that divide these approaches can be settled by resort to empirical test. Keeney and Raiffa are no doubt aware of the challenges to DT from PDT (and BDT as well), but DT is not an empirically-oriented approach, and therefore, the issues are not likely to be met head on empirically. Rather, they are more likely to be left in the realm of argumentation until the issues involving intended function, limits of competence and the relation between logic, behavioral and psychological knowledge, and practical application are examined in detail. At this point in the development of these three approaches it appears to be more useful to examine their methodological implications (see Methodology section below).

[MORE]

Integration within Group II. The principal concepts employed by Group II approaches imply different metatheories concerning the knowledge process. For example SJT metatheory describes an organism, all organisms, in their cognitive efforts to achieve stability (and thus survival) within an environment that is

"semi-erratic." Therefore, in its largest sense it lays the groundwork for a theory of cognitive evolution. It does this by providing concepts that permit the description of cognitive tasks that test the organism's ability to achieve stability. These task-descriptive concepts are not only of grand theoretical interest, however; they turn out to be of considerable practical interest because they show the ecological task conditions under which human beings (and other animals') achievement will be high or low. Thus, ecological analysis indicates that man functions well "as an intuitive statistician" when perceiving (in the correct use of the word), and much less well when making quasi-rational judgments that involve thinking as well. (The corruption of the word perception by judgment and decision researchers has not helped their scientific work.)

The same problem was addressed by Heider. He also theorized about cognitive achievement in the real world and tried to incorporate and do justice to the structural complexity, the ecology, of that world in this theory. As a result of (a) a common interest, and (b) an agreement on the environmental circumstances that create the obstacles to achievement, a general set of concepts was created that makes constructive integration a reasonable and worthwhile aim.

Much of what has been done since the publication of Brunswik's major work in 1956 and Heider's in 1958 has been directed toward the investigation of the empirical worth of these approaches. The work following from Brunswik (as represented by SJT) has on

the whole, been far more quantitative in character, and less psychological, by far, than the work that has followed from Heider. Reconciliation of those different emphases would go far toward a "third phase" that could contribute to creating a cumulative discipline of judgment and decision making.

Curiously, that "third phase" of integration might be said to have already begun with Anderson's 1974 paper. Anderson's quantitative formulation of some of Heider's hypotheses and work within the general spirit of AT is set forth in Berkowitz (1974). Yet, as noted earlier, this effort has been ignored (at least in the most recent compendium) by attribution theorists. Since the reduction of verbal formulations to quantitative ones is ordinarily rejected on metatheoretical grounds, one might suppose that Anderson's efforts toward integration have already been set aside by attribution theorists. In our view, that would be unfortunate inasmuch as previous efforts in this regard appeared to be fruitful. For example, in 1969 attribution theorists Jones, Kanouse, Kelley, Nisbett, Valins and Weiner produced a volume which contained numerous references to IIT; moreover, they actually used IIT's functional measurement in their research and employed many of IIT's principal concepts. It is strange that Anderson does not reference that work in his effort toward integration in 1974. Perhaps this sort of unevenness and distressing lack of cumulative work is to be expected until metatheoretical similarities and differences are ironed out between these approaches. It is difficult to see, for example, how AT's original emphasis (from Heider)

on environmental structure and the ambiguity it produces, are to be reconciled with IIT's indifference to this matter. But on the positive side, if the phenomenological aspects of AT can be successfully represented within the concepts of subjective scale value, weight, and cognitive algebra, then integration may proceed, with AT providing the psychological flesh for the quantitative bones of IIT.

Integration of SJT and IIT may prove to be easier precisely because SJT is more quantitative in character than AT and thus differences between SJT and IIT will be easier to detect; quantification will make empirical test easier. Metatheoretical differences may preclude such tests, for different metatheories give rise to different concepts, methods and procedures.

The metatheory of psychophysics places man in a Newtonian world in which there are (largely) independent, orthogonal physical dimensions of space, time and mass to which man reacts. Since physical dimensions are critical, psychological measurements should be made in parallel with them; hence the work on subjective scaling of physical, or more broadly, independently-defined dimensions. The metatheory of Brunswik and Heider, on the other hand, places man in a Darwinian world of flora and fauna; interdependence is the critical characteristic of this world to which man adapts, rather than reacts. Study of the single-system case is sufficient for the study of reaction; study of adaptation requires the double-system case. The quantitative (mathematical) psychologist is (usually) interested in the single-system case in which orthogonal

dimensions are emphasized, therefore, subjective scaling (i.e., of psychophysical relations) assumes considerable importance.

But the Brunswik-Heider view finds that the physicists' world (described in terms of the c.g.s. system) is not the appropriate world for the study of psychology. The distinction was made in this way by Brunswik (1956):

Many of the environmental stimulus variables mentioned by psychologists, such as "physical size" or "physical color," seem at first glance simply to be taken over from physics or chemistry. Others, such as "food," "sit-upon-ability" (William James), "likability of a person," etc., are obviously conceived with an eye to potential effects upon organisms. . . . Upon closer inspection, however, even the former often reveal psychological entanglement when they appear in the context of a psychological experiment. For example, the "sizes" of physical objects (more precisely, of physicist's objects) are in fact to be specified as "sizes of objects of attention, i.e., of potential manipulation or locomotion, of a certain human being" It is the type of organism-centered specifying redefinition mentioned above which may be summarized by saying that stimulus variables are "ecological" rather than purely "physical" or "geographic" in character.

In addition to the departure from the physicists' definition of objects and their subjective scaling, emphasis on the ecological world leads to interdependence and thus intersubstitutability of information and the ability to use such information, as is shown in the following paragraph from Hammond's 1955 paper quoted by Heider (p. 29) to indicate the congruence between Brunswik (SJT) and Heider:

Observers of the state of anger may agree that such a state exists (i.e., high reliability may be achieved), but they may not be able to communicate the basis for their decision. . . . (Hammond, 1955, p. 257). [Consider] the clinician [who] is

attempting to discover the patient's motive. The patient substitutes one form of behavior for another as he attempts to achieve his goal (equifinality). The clinician perceives these behaviors, as they substitute for one another, as cues which also substitute for one another (equipotentiality). Because of vicarious functioning, then, the clinician is hard-pressed to point at, to communicate, the basis for a decision. . . .Vicarious functioning. . . lies at the heart of the private quasi-rational nature of the clinical decision (Hammond, 1955, p. 258).

It is these metatheoretical principles that move probabilistic functionalism (and SJT) away from the single-system case and the necessity for subjective scaling taken for granted (as it should be) in the metatheory of psychophysics. Instead, probabilistic functionalism looks at survival and adaptation and those cognitive strategies taken to achieve rapport with the distal layers of the environment. Thus, for example, the Lens Model Equation (see Kaplan & Schwartz, 1975) in its most general form reads as follows:

$$r_a = G R_e R_s$$

where r_a is the correlation between the person's judgment and the empirical "answer" or response or distal variable of the environment, G is the correlation between the person's judgment and the response or distal variable of the environment corrected for the inconsistency or uncertainty or lack of control in either, R_e is a measure of the control or uncertainty in the environment and reflects the fact that for Brunswik and Heider (and SJT) the environment is an active, semi-erratic medium. R_s is a measure of cognitive control.

The important point to be observed is that this equation shows that SJT is interested in a relation (r_a) between two systems, not merely the response of one system, as is the case with Group I theories (including PDT) and IIT. And this can also be shown to be the case for AT.

Integration of Group I and Group II Approaches

There is one point of departure that may well serve to integrate these three approaches and provide a means for a cumulative effort, and that is Brunswik's lens model (see above p. II/C/3). It is so much a part of Heider's work that it might have deserved the term "Brunswik-Heider lens model." And a similar "model" has been used by Anderson (1970) and Shanteau and Phelps (1977; see Figure II-E-3). At the risk of appearing "imperialistic" we shall try to make use of the lens model as an integrative device.

A casual glance through Heider's 1958 book shows his constant use of lens model diagrams to show both vicarious mediation of information on the part of the environment and vicarious functioning (intersubstitutable use of information) on the part of the subject. In his book "An Introduction to Attribution Processes" (1975) Shaver offers the diagram presented below (Figure II-E-2) on the third page of the first chapter after the introduction.

Although Shaver's replica of the lens model is not precisely accurate, it serves well the purpose of indicating a basic platform of agreement. Shanteau's "lens model" indicates similarities, but differences as well; the reader will observe that the environmental

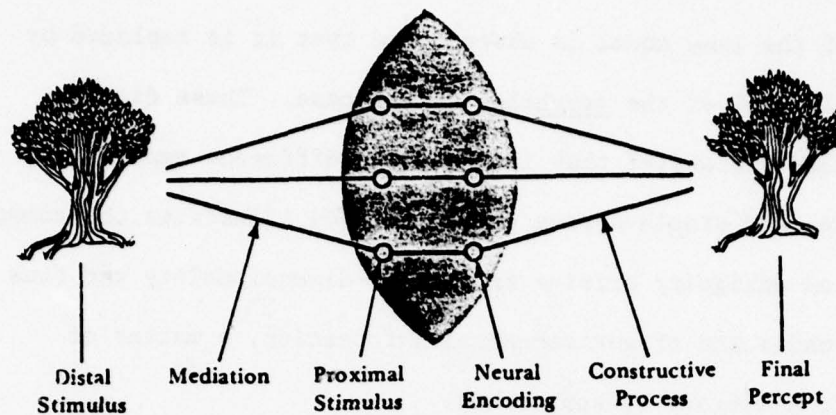


Figure 2-1. The Brunswik lens model of perception (1934) shows the course of perception from the distal stimulus in the real world to the final percept in the phenomenological world of the perceiver. (Adapted from Heider, 1958.)

Figure II-E-2

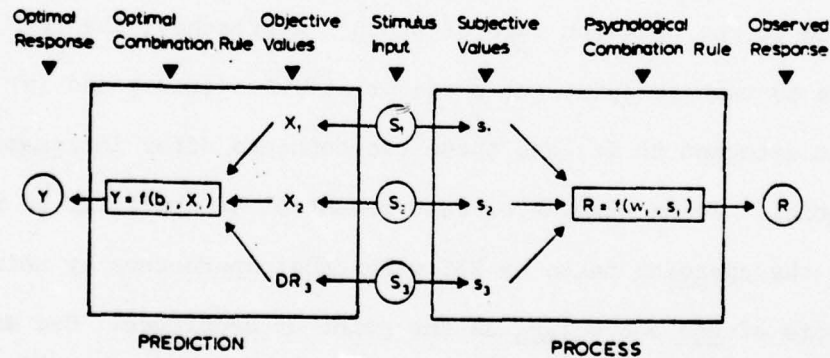


Fig. 1. General judgment diagram illustrating Prediction analyses (on the left) and Process analyses (on the right). Circled stimulus inputs (S_1) are common to both Prediction and Process. Capital letters on the Prediction side represent objective values; small letters on the Process side represent psychological values. Numeric subscripts (e.g., X_1) refer to individual dimensions, whereas alphabetic subscripts (e.g., X_1) refer to dimensions in general.

Figure II-E-3

segment of the lens model is absent, and that it is replaced by an optimal model of the psychological process. These diagrams can be taken seriously; they indicate the different emphasis in the double- and single-system case, and they illustrate the common emphasis on ambiguity arising from multi-dimensionality and thus multiple mediation of environmental information, a matter of concern to all Group II approaches.

In short, the integrative value of the lens model lies in its clear depiction of (a) the organism's effort to organize or integrate or combine or aggregate (b) various inputs or stimuli or cues or pieces of information or evidence or data, that (c) are more or less inter-related or interdependent into (d) a judgment or choice or preference or decision or impression or attribution about (e) something not immediately given.

Aside from the clarity of the depiction (easily seen to be related to the decision tree of Group I approaches) the lens model points to two concepts, (a) a cue or stimulus (etc.) and (b) the weight attached to it, and these two concepts offer integrative potential. As an example of that potential we shall try to integrate the approach taken by PDT with other approaches by using the concepts of cue and weight as the point of departure. Our aim is to classify the principal concepts that PDT introduces to explain what a probability estimate will be and to predict which heuristic will affect the judgment and make it what it is. (Our description will omit constant references to "appears to be" and "apparently"

in order to save space; the reader should understand, however, that our treatment is tentative.)

According to PDT, a person's judgment of a probability is influenced (moved up and down a probability scale) by the differential weight of various pieces of information. The general psychological principle(s) that control the manner in which these weighted pieces of information are combined is (are) not specified, therefore, we write

$$J = f(w_1)$$

simply to indicate that this aspect of other judgment theories (e.g., $J = \sum w_1 x_1$) is absent. (See also Shanteau & Phelps', 1977, comment on the absence of conceptual treatment by PDT of organizing principles.)

PDT indicates that the meaning of weight can be found in the variables that affect its magnitude. These variables are of three kinds:

1. objectively defined statistical variables (such as base-rates or [comparative] relative frequencies of the occurrence of objects or events). Results of experiments that indicate the impact of these variables (e.g., "base rates are given little weight when. . .") lead to the evaluation of the cognitive activity that leads to probability judgments. PDT generally employs problems or cognitive tasks for subjects in which hard statistical data (interpreted as probability axioms insist they be interpreted) provide "cues" for the subject who is required to infer the correct state of affairs. It is important to note that such cues have one

important characteristic: they are generally perfectly (ecologically) valid cues. In other words, they are "univocal" cues (in contrast to "equivocal" cues which generally characterize the psychological cues that confront hard statistical information, e.g., a base-rate).

Such objective, or "axiomatically safe" cues can be set in analogy to perceptual cues that are "geometrically safe" cues to size or distance (e.g., those used in the study of optics). Complete reliance on univocal cues (cues that are perfectly ecologically valid) produces perfect achievement in either case. Under ordinary circumstances, however, persons do not have access to geometrically safe perceptual cues, unless they are provided by technological aids such as range finders, etc. (That is why perception is as interesting a problem as judgment.) And, to continue the analogy, we might ask to what extent persons have access to "axiomatically safe" cues such as base-rates under ordinary circumstances. Although one can easily imagine that policy maker DMs do have access to such cues, indeed, often have them thrust upon them, PDT research strongly indicates that they do not use them (as they should) when such cues must compete with certain subjectively defined cues, of which PDT has identified one or more. (See Lichtenstein, Slovic, Fischhoff, Layman and Combs, 1978 for indications of disparities between actual rates of occurrence of lethal events and judgments of their occurrence.)

2. subjectively defined attributes of an object (such as "representativeness"). In addition to representativeness, a

further set of object-attributes has recently been introduced by PDT; thus, "the impact of the evidence depends critically on whether it is perceived as causal, diagnostic or incidental" (Tversky & Kahneman, 1977). "Specifically, we hypothesize i) that causal data has a greater impact than diagnostic data of equal informativeness, and ii) that incidental data are given little or no weight, in the presence of causal or diagnostic data." Data from the empirical analysis of the operation of these psychological variables thus provide the basis not only for (a) describing and thus (b) comparing actual probability judgments with optimal judgments, but also (c) explaining why the departures from optimality occur, and (d) predicting when they will occur.

The reader will, of course, want to raise the methodological question of the a priori status of these organism-centered definitions. We mention in passing that these are always defined in terms of plausibility arguments rather than in terms of empirical operations. That is, PDT researchers rely on the reader's willingness to agree that it is plausible to assume that a given set of circumstances is more representative than another, or that a given set of conditions implies causality whereas another set implies diagnosticity rather than carrying out experiments which provide independent evidence of subjective representativeness, etc. Few would dispute the plausibility of the assumptions made in the experiments, but it must be noted that (a) plausibility arguments do not have the status of empirical evidence, (b) the (somewhat indistinct) road of advances in psychology is littered with the

wreckage of plausibility arguments, and (c) that it may well be difficult to ascertain organism-centered definitions of these variables in the more ambiguous circumstances of real world (test) questions.

Be that as it may, as indicated above, few would (and indeed, no one has) dispute(d) the plausibility of the operational definitions of these organism-centered variables in any specific case, nor have the psychological definitions received any criticism; a matter to which we turn in a moment.

Before turning to the third class of variables, we should note that the research method used by PDT is similar to that used by Group II approaches, namely, to "confront" various types of cues with one another in order to discover their relative weights under various conditions. PDT frequently pits objective cues such as base rates against subjective cues such as representativeness. That is, the task is presented in such a way that it offers the subjects a choice between "axiomatically safe" cues (e.g., base rates) and "psychologically safe" cues (such as personality data) [Note: the terms "axiomatically safe" and "psychologically safe" are introduced by the authors, not PDT researchers].

The general results of these "confrontation" studies that are emphasized by PDT indicate that "axiomatically safe" cues are ignored in favor of "psychologically safe" cues. Recent efforts ("causal schemata"), however, are more specific and indicate under which conditions each type of cue will carry how much weight.

Confrontation studies of this sort have a long history in psychology, particularly in perception, where various types of cues are pitted against one another. (See Brunswik's, 1955, use of the term "diacritical confrontation.") And the effort toward cumulativeness may well be served by translating the term "heuristic" (valuable as it may be) in this way. The danger in such a translation is clear; a heuristic is a device, a "rule of thumb," thus it implies more than a cue--which is a "hint." A heuristic implies a strategy, a cue implies an "input." Whether PDT has developed concepts that refer to strategies or cues is considered in more detail below. Here we turn to a different type of heuristic.

3. types of cognitive processes (such as "availability" or "anchoring"). Data from the empirical analysis of the operation of these psychological variables also provide the four functions of (a) describing, (b) comparing, (c) explaining, and (d) predicting (indicated above).

Separating these three classes of variables is not easy. Although objective denotation of statistical variables is simple enough, and therefore easy to separate from the two subjectively defined classes of variables, separation of the latter is not so simple, yet separation does seem to be required. In particular, separating the denotation of attributes from process variables is not as easy as one might hope, although the attempt at separation was made by PDT at the outset. Thus, availability was described at the time of its introduction (1973) as "associative

distance," to be contrasted with representativeness, which was described as "connotative distance" (p. 208): "one estimates probability by assessing similarity or connotative distance. Alternatively, one may estimate probability by assessing availability or associative distance." That distinction indicates that we have two types of "distance," and the distance appears to apply to objects, that could perhaps be scaled in these terms. Thus, we are dealing with attributes. But we could not find this distinction appearing again. There are, however, numerous references to representativeness as a quality or attribute assigned by the person to independently real objects or processes: Example: "Which is more likely? This series (HTTHTH) or this series (HHHTTT)?" Persons say the former because it subjectively 'best represents both the population proportion (1/2) and the randomness of the process' (plausible reasoning). That is, representativeness is an attribute of these processes and can be used as a cue when inferring the likelihood of the occurrence of either process.

Availability, on the other hand, is described as the ease with which the relevant mental operation of retrieval, construction or association can be carried out (1973, p. 208). Availability therefore refers to the process of recall: some objects or events are more available, however, because of their special characteristics, notably their "associative bonds." "The availability heuristic exploits the inverse form of this law, that is, it uses the strength of association as a basis for the judgment of frequency. In this theory, availability is a mediating variable,

rather than a dependent variable (see above) as is typically the case in the study of memory" (p. 209). Here we encounter difficulty; availability refers to the ease with which a "mental operation. . . can be carried out in memory" as a result of the association between objects (or processes). But the associative distance between objects can also be thought of as an attribute of an object or objects, and thus can be considered to be a cue. Indeed, Tversky and Kahneman refer to it in just those terms: "Availability is an ecologically valid clue for the judgment of frequency because, in general, frequent events are easier to recall or *imagine than infrequent ones*" (italics ours, p. 209). Although this definition was later changed to read: "availability is a useful clue because instances of large classes are recalled better," the point remains. In this usage availability refers to an attribute of independently real objects, as well as to the process of recall itself.

The uncertainty surrounding the conceptual status of availability can be seen more clearly by using the concept of "anchoring" as an anchor. Anchoring is clearly a psychological process; it is not an attribute of an object (as is representativeness).

There are other points at which integration might be achieved through the use of the "lens model analogy" that go beyond the concepts of cue and weights attached to them. Thus, for example, a central concept in SJT is "cognitive control"; it refers to the control persons can exercise in the execution of the knowledge they may possess (a concept that has turned out to be of particular

value in Gillis' application of SJT to the judgment processes of schizophrenics and the differential effects of various psychoactive drugs; see Hammond & Joyce, 1975). SJT explains variations in cognitive control as due largely to environmental circumstances (mainly uncertainty). But Tversky's treatment of choice as elimination by aspects (1972) includes the following paragraph which suggests a different (internal) source of decreased cognitive control.

The EBA model accounts for choice in terms of a covert elimination process based on sequential selection of aspects. Any such sequence of aspects can be regarded as a particular state of mind which leads to a unique choice. In light of this interpretation, the choice mechanism at any given moment in time is entirely deterministic; the probabilities merely reflect the fact that at different moments in time different states of mind (leading to different choices) may prevail. According to the present theory, choice probability is an increasing function of the values of the relevant aspects. Indeed, the elimination-by-aspects model is compensatory in nature despite the fact that at any given instant in time, the choice is assumed to follow a conjunctive (or a lexicographic) strategy. Thus, the present model is compensatory "globally" with respect to choice probability but not "locally" with regard to any particular state of mind.

Whether the use of the lens model as an integrative device will be found to be acceptable remains to be seen, and whether the lens model analogy can be usefully pushed further into the domain of Group I approaches than PDT, also remains to be seen. We shall point to that possibility in our discussion of the loci of the concepts employed by various approaches.

F. Loci of Concepts

Introduction

In this section, the concepts employed by judgment and decision making theorists are described in terms of their location within a reduced form of Brunswik's "lens model" diagram. Our use of the lens model diagram does not, of course, imply that the reader must become immersed in, or committed to, the Brunswikian framework. We use the diagram simply because it is broad enough to accommodate the concepts employed by all six approaches, and thus serves as a useful device for locating redundancies as well as gaps in our theoretical efforts. Thus, for example, the Figure II-F-1 indicates seven areas of the judgment and/or decision process to which theoretical concepts can be applied.

In what follows we shall attempt to indicate the extent to which the six approaches have developed concepts (or have assigned theoretical importance to) each of the seven areas indicated in Table II-F-1.

Table II-F-1

1. Refers to problem of selecting the cues that enter into a judgment
2. Refers to problems of subjective measurement and intuitive evaluation of data; turning "soft" data into "hard" data
3. Refers to problems of subjective use of objective, "hard" data; scaling
4. Refers to problems involving ecological and subjective inter-correlations among cues

Figure II-F-1

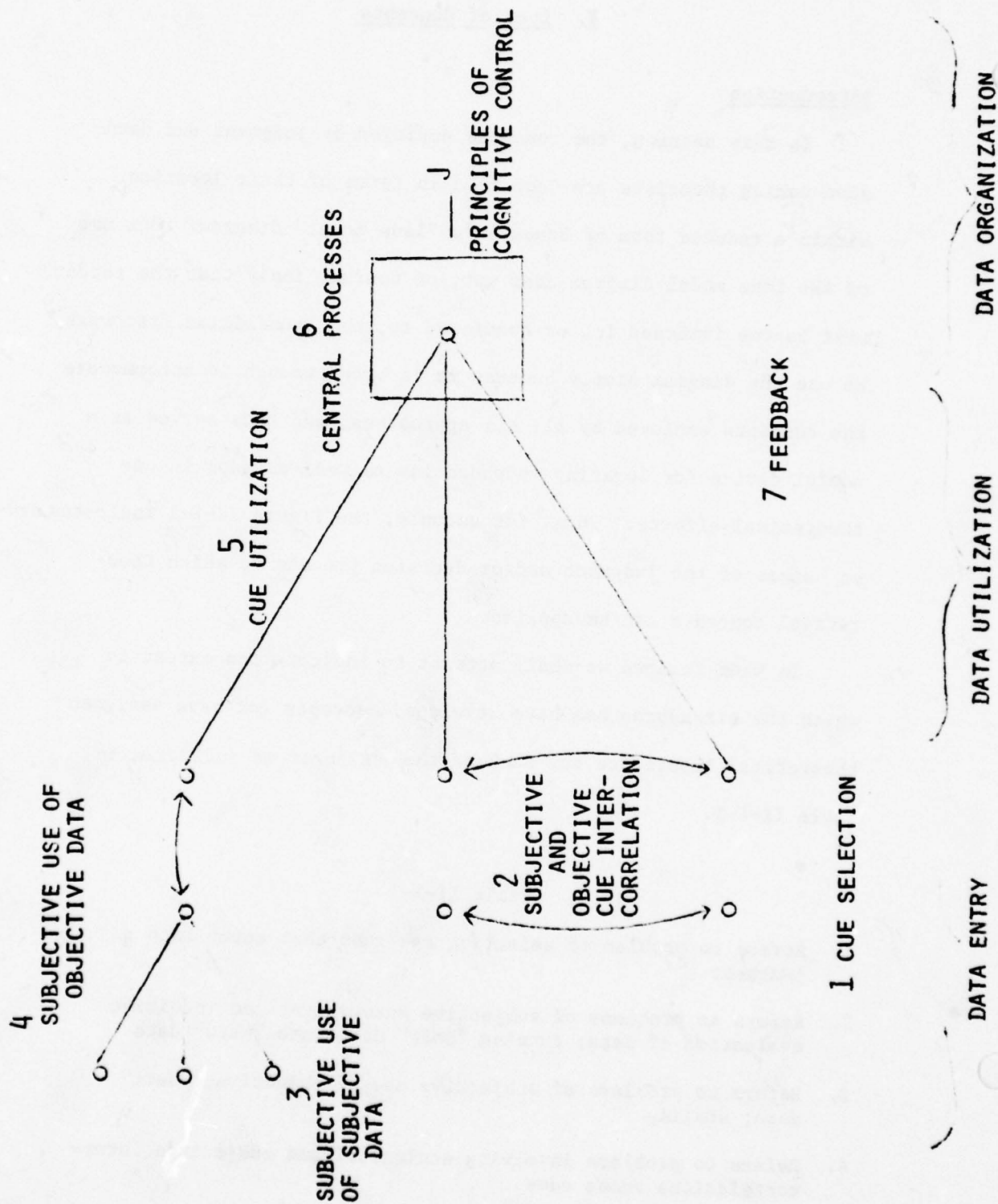


Table II-F-1 continued.

5. Refers to problems involving subjective weights; function forms
6. Refers to problems involving principles of organization, e.g., additive or configural principles, cognitive control
7. Refers to problems involving feedback.

Decision Theory

This approach locates its concepts centrally; there are no theoretical concepts that refer to what is now called the "processing" of information prior to the person's expression of a subjective probability. Judgments of probability are the result of prior knowledge gained from experience, education, theory, etc., prior to the evocation of the probability judgment. For example, when the decision maker offers a subjective probability (p) that an event (e) will occur, s/he is offering a complex judgment based on information that is somehow organized into that judgment. Decision theorists, are, of course, prepared to decompose, indeed, to urge the decomposition of "larger" subjective probabilities (and utilities) into "smaller" ones by means of decision trees, that is, to "divide and conquer." But there is no theoretical attention given to the process by which past or present information, or data, is transformed into subjective probabilities, nor to the process by which such division is made. In short, processes or items 1 through 5 are not given theoretical attention. Careful procedural attention is given, however, to the matter of obtaining and measuring, i.e., calibrating, the decision maker's probabilities and utilities based

on the acceptance of a metatheory regarding psychological measurement shared with BDT, PDT and IIT (but not SJT or AT; see below).

Such procedures (e.g., the use of lotteries) are justified on either (a) pragmatic grounds ("this works better than that, at least in my experience") or (b) theoretical-logical assumptions about measurement. There is no explicit theoretical concern for the objective world of event-frequencies and utilities and the process by which these give rise to subjective probabilities.

Finally, the decision theorist "knows" that the central cognitive process of aggregation of probabilities and utilities is nearly always nonoptimal by any prescriptive criterion. Without considering the wide variety of possible forms that psychologists have imagined cognitive activity might take, the decision theorist prescribes the optimal form and substitutes it for whatever actual form (or forms) it (or they) might be, and thus "aggregates" the probabilities and utilities for the decision maker.

In short, all the major concepts employed by the decision theorist are located centrally. The activities of decision theorists are directed toward developing measures of these central concepts, with little or no regard for psychological research that might bear on proximal-central relations or procedures used to measure them.

Research, therefore, is logical-theoretical rather than theoretical-empirical; it is devoted to the development of the mathematics of the prescriptions for various types of problems. Empirical work is given over to applications that consist of case histories indicating the usefulness of certain steps (e.g., divide

and conquer) toward rational decision making. Although the extent of that work is impressive (see Keeney & Raiffa, 1976), it is not evenly distributed over the major concepts. As Keeney and Raiffa note in their Preface (p. viii) DT has given far more attention to the "uncertainty phase" (that is, the ascertaining of probabilities and the development of axioms of probability theory) than to the "value or preference side of the picture." Both concepts have received considerable theoretical attention elsewhere, however; probability mainly from mathematical statisticians, utility mainly from economists. We do not include these theoretical efforts as part of the present discussion, however, because the theoretical differentiations in that literature have little relevance to the present work of DT.

Behavioral Decision Theory

The work carried out within this approach is best understood if research and application are separated.

Research. Many of the concepts employed by researchers in this approach flow from the Bayesian model of organizing data concerning events into a probability judgment about the truth or falsity of hypotheses. As a result, the research includes concepts that refer to (a) events on the periphery of the organism (e.g., a "datum"), (b) concepts that refer to the translation of external data into a judgment of a probability regarding the likelihood of the truth of a hypothesis, and, of course, (c) central concepts of judgments of subjective probability and (d) their revision (with or without

feedback). Bayesian researchers construct decision tasks and manipulate the parameters of such tasks (in contrast to DT researchers), and therefore they employ a set of concepts to describe the task circumstances. These include probabilities, events, situations (e.g., "stationarity"), and relations between events (e.g., dependence) as noted above. In particular, Bayesian research includes the concept of "optimality" with regard to the subject's behavior; "optimality" is evaluated in terms of the extent to which the subject's (posterior) probability judgments about the truth of a hypothesis approach those produced by the Bayesian model provided with the same data. Because (a) the parameters of a task are specified and manipulated, and (b) a criterion regarding the achievement (Bayesian optimality) of a subject is provided, Bayesian research may be said to be conducted within the framework of the "double-system" case (the task constituting one system, the subject constituting the other). Therefore, insofar as Bayesian research is concerned, the above single-system lens model must be extended to include the "second system" that represents the task as in the original lens model diagram. Thus Bayesian research studies learning and examines feedback. (See relation to SJT with regard to cognitive feedback.)

There are (a) distal (causal) categories ($H_1, H_2 \dots H_n$), (b) objective descriptions of cues (e.g., chips) (proximal stimuli), and (c) there can be conditional dependence, etc. between cues, and (d) cue utilization can be observed in the "impact of a datum"/per event, as well as (e) organizational principles imputed (as a

hypothesis) to the subject, such principles in the form of a "likelihood ratio" or Bayes' theorem. And it is the comparison between the Bayesian "likelihood ratio" (or some other aspect of Bayes' theorem) and the subject's likelihood ratio (or other response) that provides a measure of the subject's achievement as an "intuitive statistician" or adaptive cognizer under conditions of uncertainty.

Thus, it appears that the research paradigm employed by BDT and Bayesians regarding the "extraction of information" or regarding the accuracy of the estimation of posterior probabilities, or etc., is one which explicitly employs theoretical concepts that apply to all seven areas of the diagram. It should be noted, however, that all of the concepts converge on a single end term (or response) namely the expression of a probability judgment.

[MORE on cascaded inferences]

Application. The distribution of theoretical concepts in the application of Bayesian theory is not as clear as in the research situation because application includes MAUT. First, we observe that MAUT has been applied to two sets of task circumstances (a) the n-system case in which the multiple systems are several decision makers without a known task system (in contrast to the situation described above in which the two systems include one decision maker and one task) and (b) the single-system case in which there exists only a decision maker, the task parameters being unknown. We discuss the n-system case first.

Here we follow Gardiner and Edwards (1975) in their description of the application of MAUT to social decision making in which

it is assumed that interpersonal and intergroup conflict will occur. Leaving no doubt about the importance of this application the authors say (p. 35) that "We consider the idea of resolving value conflicts at the level of decision rules (rather than at that level of individual decisions) to have the potential of revolutionary impact on land-use management--and many other public-decision contexts as well."

The outline of the process of application of MAUT indicates four main phases.

Phase 1. Definition of the nature of the problem and its dimensions. This step is primarily substantive. It refers primarily to the specification of the attributes (cues, in Figure II-F-1) that the decision makers want to consider; item 1 in Figure II-F-1.

Phase 2. Probability evaluation. This step is one of extracting a probability judgment. Gardiner and Edwards say (p. 10) that "other names used for the same process in other contexts are diagnosis, intelligence evaluation, data gathering and so on." But the "output (i.e., response or judgment) is a set of posterior probabilities of states" in all cases. They also indicate that "less formal judgments of probability. . .and such formal ones as the acceptance or rejection of statistical hypotheses, also fit here and might be substituted for the Bayesian version of the process."

Probability evaluation is evidently a cognitive process that intervenes between (a) the subject's observation of the several attributes of various objects and (b) the subject's response that indicates the probability of occurrence of various "states."

Thus if Step 1 refers to dimensions (cues, attributes; item 1 of Figure II-F-1), Step 2 refers to the judgment (item 6 of Figure II-F-1) of the probabilities of occurrence of the states so described.

Phase 3. Outcome evaluation (SMART). This step refers to "the attachment of values. . .to outcomes." Part of what makes the process difficult is the fact that outcomes include "a multiplicity of value dimensions. . ." which "presents a multiplicity of problems" (p. 13). Thus, the process of outcome evaluation parallels the process of probability evaluation. (See also SJT). Gardiner and Edwards make it clear that without the assistance of a MAUT decision analyst, ordinary decision makers confuse these two processes; a decision analyst separates these processes, however, and combines them analytically and thus produces a rational choice for the decision maker.

As in probability evaluation, a set of attributes (dimensions) characterize each of a number of outcomes. And as in the case of probability evaluation, direct estimates (of amounts of probability in the case of probability evaluation, of amounts of importance in the case of outcome evaluation) are obtained simply by asking the decision maker for them (in contrast to DT's more complex method of requiring the decision maker to choose between lotteries, or SJT's method or IIT's method of extracting them post hoc).

Interdependence between cues. (Item 2, Figure II-F-1). Interdependence of cues or data is of considerable theoretical importance to MAUT. Two types of interdependence are acknowledged within the

SMART approach (interdependence is given detailed analysis in Bayesian research by Edwards and his colleagues). One type of interdependence is that of value dependence, the other is called environmental interdependence. (Note: Gardiner and Edwards call this "value-independence" and "environmental-interdependence"; p. 20.) Since values are entirely subjective and environmental facts are objective (i.e., measured independently of the decision maker) this distinction can be represented in Figure II-F-1 as the distinction between objective and subjective intercorrelations. As Gardiner and Edwards observe (p. 20) these may not match.

Therefore, MAUT (in the form of SMART) approaches the problem of ascertaining (or externalizing) utilities or values precisely as it does probabilities. That is, from a theoretical point of view, in both cases it specifies cues, their interdependencies, their differential utilization and their organization.

Psychological Decision Theory

Because the PDT approach takes the SEU theory as its point of departure, its fundamental concepts are probability, utility, and (although seldom mentioned) aggregation. But the PDT approach has argued that the empirical evidence indicates that the departures from optimality begin with systematic errors in estimating probabilities (including randomness) and that these systematic errors are explained by psychological concepts such as availability, representativeness and anchoring (see Principal Concepts for definitions).

As noted above, all of these concepts are explanatory, rather than descriptive. They explain why some cues are over- or under-utilized, and thus explain why judgments are often found to be in error. These concepts are directed toward the formal aspects of the judgment task as well as substantive aspects of the task. That is, Kahneman and Tversky have pointed to situations where the statistical properties of the judgment task themselves give rise to errors of judgment because of their (un)representative character, as well as situations where the substantive materials of the task give rise to errors of judgment because of their (un)representative character.

Because of their explanatory character, these concepts cover three major aspects of the judgment process: (a) cue-utilization, (b) weight, and (c) central processes. That is, availability, the central activity, explains why certain cues are weighted more or less heavily in any specific situation. (Approaches that are largely descriptive do not make such links.)

Areas of the lens model diagram (Figure II-F-1) that are affected by these explanatory concepts are (a) cue selection, (b) subjective use of subjective data, (c) subjective use of objective data, as well as (d) cue-utilization (weight) and (e) central processes. It should be mentioned that although these concepts clearly refer to central processes they do not address the matter of organizing principles--the form of aggregating or integrating data into a judgment. No examination of

learning has been undertaken by PDT and it has therefore not addressed the problem of feedback.

[MORE]

Transition

Theoretical concepts within the first three approaches to which the probability-utility-aggregation premise is basic can be found at all critical points of the lens model representation of the judgment process. There are large differences among the three approaches, however. DT offers the least coverage, whereas BDT (Edwards) provides concepts for each of the seven processes considered in the diagram. Although the PDT approach does not treat each of these points as explicitly as BDT, its explanatory efforts add to the focus of the coverage it provides.

Social Judgment Theory

This approach includes theoretical concepts that apply to all seven areas of the lens model diagram (as in the case of BDT). Second order concepts (e.g., false agreement) also appear in connection with research on conflict.

One point deserves special mention, however, and that relates to Item 3 on the lens model diagram, namely subjective use of objective data, that is, the scaling problem of traditional psychophysics. SJT does not treat this problem theoretically in the traditional way. (As will be seen, IIT gives careful attention to this problem; see Principal Concepts, Integration Section for further discussion.)

[MORE on this]

In addition to applying theoretical concepts to the seven areas of the single-system case, SJT includes several overall concepts that apply to the double-system case (e.g., zone of ambiguity, causal texture, etc.) and gives special emphasis to the concept of feedback with regard to learning to improve judgments (as does MAUT) and with regard to the difficulty if not impossibility of improving achievement in the absence of such feedback.

SJT also has developed theoretical concepts that apply to the processes of interpersonal learning (both from and about others) as well as to interpersonal conflict.

Information Integration Theory

This approach has by and large restricted its work to the single-system case although there have been demonstrations of its ready applicability to learning and to group judgment making. Primary emphasis has been given to the concepts of (a) valuation (item 4), (b) weight (item 5) and (c) integration (organizing principle) (item 6), although all other aspects of the lens model diagram have been dealt with in some form (e.g., item 2, subjective intercorrelation among cues, has been treated in terms of "discounting"). And within these three major concepts, the organizing principle, or integration process, has received greatest emphasis. Within the framework of the single-system case depicted in Figure II-C-1, then, theoretical concepts applying to six areas can be found. The major omission is that of feedback; no theoretical

analysis of the concept has appeared as yet, presumably because the major emphasis lies in the study of the single-system case.

Specific and detailed treatment is given to peripheral, mediating, and central phenomena. IIT employs a variety of concepts concerning peripheral (i.e., stimulus) data, some from psychophysics (e.g., contrast), some from social psychology (e.g., change of meaning).

Treatment of mediating processes ("cue" utilization) is given less theoretical treatment but it is certainly there in principle in the concept of weight.

Central processes are described wholly in terms of organizing or integrating functions that can be described algebraically (e.g., data are added, averaged, or multiplied).

[MORE]

Attribution Theory

It is easy enough to document the assertion that Heider's approach includes concepts that apply to all seven areas of the single-system case. In addition, this approach also includes, and gives detailed treatment to, one aspect of the judgment and decision problem none of the other approaches do and that is the effect of the subjects' action on his attributions, inferences, judgments and decisions, as well as the reverse. But little or nothing is said about weight, an omission that very likely follows from the lack of quantitative emphasis. Note that when Anderson's functional measurement is applied to AT (see above) the concept of weight is readily employed.

Integration

If a lens model diagram were to be constructed showing the loci of all the theoretical concepts employed by the six approaches considered here, it would indicate that all seven areas in the diagram would be covered by concepts from at least two approaches, with the exception of the n-system interpersonal conflict situation, and even here, the scarcity of concepts is more a matter of benign rather than deliberate or antagonistic indifference. (Indeed, as Mumpower [1976], has shown there is considerable overlap between AT and SJT in the case of interpersonal learning.) Therefore, if one assumes that the lens model diagram (Figure II-F-1) does in fact adequately depict the major aspects of the judgment and/or decision making process, then it may be said that all aspects of this process have received, at the least, *some description, and on occasion, some explanation*. Redundancies, rather than omissions, characterize the situation.

The next question, of course, is: Can these heterogeneous concepts be "integrated," or somehow merged into a single, semantically unified set of concepts, or are there antinomies among them that must be settled either by further analysis of their meaning or empirical test.

An example of the unification is this: BDT researchers recognize and give due importance to the concept of interdependence among attributes. SJT researchers also emphasize the importance of what they call the "ecological intercorrelation" among cues. Deciding that the Bayesians "conditional dependence" should be

applied to the interdependence among categorical or nominal attributes, while "ecological intercorrelations" should be applied to the interdependence among objective cues or variables and recognizing that both terms refer to the contingencies or covariation between objective variables would be a simple matter of convergence. (The same would hold for AT.)

Antinomies are another matter. For example, when Kahneman and Tversky (1972, p. 450) assert that ". . . man is apparently not a conservative Bayesian: he is not a Bayesian at all" we have an antinomy that must be resolved analytically or empirically.

But this is not an antimony that extends beyond a contradiction in explanatory or descriptive statements of theoretical content. We are not observing a theoretical dispute about whether some aspect of the judgment process should be or must be eliminated. Such a situation does not seem likely to develop in the near future.

That conclusion may result from our present use of a coarse theoretical framework, however. Future analyses that use a finer method of description and comparison may find such intractable antinomies regarding scope. At present, however, we conclude that (pride of authorship aside) it would not be difficult to move toward a cumulative discipline by developing a semantically unified theoretical framework.

G. Intended Uses of Research Results

Introduction

It may appear strange to include this topic in a discussion of theory. But all of the Group I approaches and at least one of the Group II approaches (SJT; possibly IIT also, see Shanteau & Phelps, 1977) have made it clear that they intend their work to be used not only by other scientists but by policy makers, and, indeed, by anyone who has to make a decision of more than trivial consequence. Therefore, we need to discover to what extent these approaches have been directed toward judgment and decision making in the real world as against the laboratory. For example, it is generally acknowledged that in the real world, task variables (as they are ordinarily described by the judgment maker or decision maker) are interdependent. But including such interdependencies in empirical research creates severe problems of interpretation and it is seldom done. One question that can be asked, therefore, is to what extent such interdependency is included within the theoretical framework of those approaches that intend to apply their work to the real world. A second question that follows from the first is: what theoretical guidance is provided for dealing with, for example, interdependencies among task variables? Put otherwise, how does theory guide method in application, if, indeed, it guides it at all? In addition, all but two approaches (IIT, AT) have indicated their interest in developing "aids" for the policy maker and/or decision maker. To

what extent and in what way do the concepts included in these theories direct the construction of such "aids?" And to what extent are judgment aids different from decision aids? These questions cannot be answered properly in this report but they deserve an answer; we need to know whether theory and the results of basic research control application, whether application proceeds independently, or whether, in fact, the distinction between basic research and application can be made at all in the study of judgment and decision making.

Decision Theory

The mathematical research that elaborates and extends the logical entailments of SEU theory is intended to be used to aid decision makers to achieve a better procedure for reaching decisions. "Better," of course, is defined in terms of the approximation of the procedure to the axioms of DT. It follows that if the decision procedure can be justified in terms of the mathematical logic of DT, and if it is followed, the decisions themselves will be better. In short, the use of DT, decision theorists argue, is entirely justified by the logical consistency of the mathematics that supports the prescriptions it offers to decision makers.

Although it is obvious that DT is thus prepared to defend the objectives it seeks, it is uncertain whether DT would argue that it does not matter much how logical consistency is achieved. Is there, for example, a proper amount of persuasion and an improper amount? If so, how much persuasion is improper? But it is almost certainly

true that DT would argue that the application of DT should be postponed until the research is done that will tell us whether indeed it does matter how logical consistency is achieved. DT's primary consideration seems to be directed toward reducing the likelihood that the probabilities and utilities are those of the decision analyst rather than the decision maker. And that seems to be entirely a matter of personal technique.

Operational questions aside, however, it is clear that the theory that supports the aims (not the specific techniques) of decision analysis does have a strong directive influence; it certainly tells the decision analyst what information s/he might get under specific circumstances and, very often, what approximations are satisfactory under specific circumstances.

Behavioral Decision Theory

BDT theorists see two uses for the results of their research: (a) descriptions and explanations of decision making behavior that will find a place in the basic science section of psychology, and (b) the development of procedures (decision aids) that will enable decision makers to make better (i.e., logically defensible) decisions. It is only with regard to (a) that the efforts of the behavioral decision theorists differ significantly from the efforts of the decision theorists.

Basic research uses. BDT efforts are generally directed toward "do people do this or do that?" questions. Do people aggregate probabilities and utilities by adding or multiplying them? To what extent do people organize information and act on it cognitively

(intuitively) as if they were Bayesian statisticians? In general, departure from the Bayesian model are of central interest.

Application. The results of Bayesian research have been applied in a variety of decision making situations (see, for example, Edwards, 1972; Fryback, 1974; and Edwards, 1978). Present uses of BDT involve SMART (see Gardiner & Edwards in Kaplan & Schwartz, 1975). Bayesian research is hardly mentioned in that application, although MAUT is described in detail. The same is generally true with regard to the use of SMART in evaluation research; the results of Bayesian research do not bear directly on the use of SMART. In Gardiner and Edwards (1975), for example, there is little reference to research regarding the number of dimensions employed, how weights are to be ascertained, methods of coping with interdependencies, etc. (points which are discussed in detail in the Procedures section). Nor is there any reference to any research dealing with interpersonal conflict.

In short, although BDT differs from DT in that it carries out empirical behavioral research, the results of that research are not often brought directly to bear on efforts to aid the decision maker. Perhaps this omission is more apparent than real, however, and reflects the style of the presentation.

Psychological Decision Theory

Until Kahneman and Tversky's article in Management Science (in press), it could be fairly said that although PDT had pointed to the fallibilities rooted in human judgment and decision making, PDT had

not indicated what remedies should be applied. Remedies are indicated in the Management Science article, however, and it is clear that these are developed directly in response to the "biases" that these authors and others have found in their research. In short, having found certain biases, they offer "debiasing" procedures.

These are entirely statistical; for example, "To correct for nonregressiveness, the intuitive estimate should be adjusted toward the average of the reference class." Thus, PDT offers logical mechanics to guard against the use of psychological mechanisms.

As noted earlier, however, PDT has yet to develop theory regarding central, organizing principles and does not address the problem of interdependencies among "inputs." Since no mechanisms for dealing with these are offered, it is probably fair to say that theory and assistance proceed in parallel fashion.

There is no empirical proof of the utility of the remedies offered by Kahneman and Tversky in their Management Science article and in one sense none is needed; the statistical logic offered has a better claim to truth than the "psychological logic." In another sense, however, proof is needed as it is in the case of other decision aids. Can it be used? Will it be used?

Lichtenstein, Slovic, Fischhoff, Layman, and Combs (1978) attempted to debias their subjects without success. In this case, the debiasing method was used; it simply didn't work. But since this was an initial attempt, one can only conclude that debiasing that is not simply application of statistical techniques may require hard work. Be that as it may, it is clear that efforts to debias

decision makers follow directly from the results of basic research as well as the logic of DT.

Transition

The conceptual distinction between judgment and decision that was made above (see Intended Functions of the Theories) can be seen in its operational form in the differences in the construction of "decision aids" within Group I and "cognitive or judgments aids" constructed within Group II.

Social Judgment Theory

As anticipated, Group I provides decision makers with information about their subjective probabilities and utilities (a) in the "divide and conquer" approach to a decision tree (DT), (b) in SMART or MAUT (BDT), and (c) in debiasing (PDT). Group II judgment analysts (represented only by SJT), on the other hand, provide policy makers with information about relations between their judgments and cues, as well as information about other persons' judgments in relation to task properties. SJT insists upon such cognitive feedback because such relations are argued to be the basic unit of knowing; indeed, SJT insists that such cognitive feedback is the only means of inductive learning, inside and outside the laboratory. Thus it might be concluded that decision aids that deal with probabilities and utilities address themselves to the results of inductive knowing, that is, cognitive state of an individual subsequent to the acquisition of knowledge, whereas judgment aids address

themselves to the sources of those probabilities and utilities, such sources being described in terms other than probabilities and utilities.

As noted earlier SJT derives its arguments for its form of an aid for judgment from the results of theory and research regarding the double and n-system cases. The clearest examples of the insistence of the SJT theorist upon environmental representation in judgment research can be seen in the recent development of judgment aids that link environmental models with models of judgment systems (see, for example, Hammond, Mumpower, & Smith, 1977; Hammond, Klitz, & Cook, 1977).

These two worked-out examples are intended to show that neither set of modelers (environmental systems modelers or cognitive modelers) can work effectively in isolation, and that the policy maker must have access to both models and their interactions if s/he is to be successful. This argument, of course, follows directly from Brunswik's admonitions that

Both organism and environment will have to be seen as systems, each with properties of its own. . . Each has surface and depth, or overt and covert regions. . . . It follows that, much as psychology must be concerned with the texture of the organism. . . it must also be concerned with the texture of the environment (1957, p. 5.).

Until the advent of the computer model of the environment this admonition was honored more in its breach than in its execution, mainly because of the technical difficulties of bringing the environment under control. Computer models can effectively solve that problem, however, and can bring the interaction between policy

makers and the complexities of their environment under control. It is in this way that SJT follows the above metatheory regarding organisms and environment in its effort to aid the policy maker.

Information Integration Theory

This approach has generally stood aside from the matter of application. The only example that we know of is that provided by Phelps and Shanteau (1978) regarding the judging of animals. Even in this case however, the investigations seemed more interested in testing the descriptive accuracy of various models than in providing a judgment or decision aid.

Attribution Theory

Although AT clearly intends that its research results are used to develop general laws of inductive(?) knowing or inferences about the locus of causality, and although there are no direct efforts to apply the results of AT research to specific problems in the real world, nor are there attempts to develop "attribution aids," there is an implicit suggestion that the results of AT research can be (or at least will be) directly applicable to behavior in the real world. Indeed, because of its social, rather than physical, character, the empirical work in AT implies that the results grow out of or are based on simulations of the real world and are thus directly generalizable to it. The realization of these aims seems to be far in the future, however.

Integration

The intended uses of the research results produced by decision analysis (Group I) and judgment analysis (Group II), although different, are not highly divergent. Group I intends its theory to be useful to decision makers in general and policy makers in particular; SJT, DT, BDT, and PDT also clearly ask their theory to be useful, to guide them toward producing results that not only contribute to knowledge about the psychology of information processing, but also contribute toward the development of aids for policy makers. All four approaches want to be able to point to theory-based, research-based, applications. But there are major differences between and within and between these four approaches that follow from their theoretical (and metatheoretical) orientations. To what extent can these differences be resolved and/or set aside in the development of aids to the policy maker?

We offer only one suggestion: if judgments and decision processes can be usefully separated with regard to their intended theoretical function (as described above) then they may be usefully separated with regard to their pragmatic function. That is, Group I approaches may aid the policy maker to explicate his probabilities and utilities and to organize these in a logical manner, thus avoiding the biases and fallibilities of unaided cognitive efforts, whereas Group II approaches may aid the policy maker to see sources of these probabilities and utilities, and to see the relations between various judgment policies and their implementation in a

world simulated by models that vary in their "hardness," or analytical strength.

The combination of these two approaches to providing aids for the policy maker (or decider) demands exploration, theoretically and methodologically, as well as empirically. The creation of a useful judgment and decision environment for policy makers and other deciders deserves immediate attention.

[MORE!]

Addendum to "Intended Uses of Research Results"Claims for the Value of Various Approaches

In the course of our study of the different approaches, we frequently encountered claims of benefits on behalf of those approaches. This was particularly true for those approaches that have been used outside the laboratory. We have collected and categorized many of those claims in the belief that such claims can serve as useful criteria for evaluating and comparing the different methods and procedures, although we have not yet undertaken that task, however. At the very least, these claims indicate important psychological issues which need further exploration in the context of judgment and decision making. Almost all proponents, for example, claim that their respective methods are useful for teaching novices how to arrive at better judgments, or how to resolve group conflict about a decision. The most striking feature of the list of claims is that seldom have these claims been verified experimentally; rather they rely on references to case histories.

The categories in the list below are not always mutually exclusive (that may be impossible and undesirable anyway) and the list is undoubtedly incomplete. We have only included those claims seen in print; often a direct quotation making the claim is presented. Because the claims of the approaches (at least for DT, BDT, and SJT) are so similar and because we would be surprised if any approach would deny that its method could produce such benefits,

we do not indicate the sources of the quotations; representatives from each approach are encouraged to try to identify their own claims--as well as others!

Claim 1: Aid to clear thinking

Almost all claims fall into this category, but not everyone is explicit in claiming that a decision maker's (DM from now on) thinking will be improved. Examples: "a framework for straightening out one's mind," "creating cognitive aids for human judgment," "extending the limits of human judgment."

Claim 2: Educates the DM

Whether or not these methods help DMs think better, almost everyone claims that such methods make the DM wiser about his particular problem. "There is educational value in asking more specific questions, getting specific utility functions, and so on. It sensitizes the decision maker. . ." Below is a list of more specific educational claims.

- a. Makes hidden assumptions and implicit tradeoffs explicit
". . .we should not underestimate the important intellectual and emotional impact that arises when we are forced to express vexing tradeoffs in unambiguous terms." ". . .help management articulate some of its basic assumptions."
- b. Forces full consideration of all consequences of a given action.

Why conduct an experiment or do an opinion survey if there is virtually no expectation that action will be altered by the

information obtained? "Explicit definition of the possible outcomes tends to prevent this common trap."

c. Identifies what is important for making decision and where more information is needed.

"Examination of value of information in a decision context helps suggest the gathering, compilation, and organization of data from new sources." DM "can come to grips with the critical determinants of his decision." Causes DM to "grapple with fundamental issues."

d. Identifies what is not important for making the decision.

This claim is obviously a by-product of (2c). DM can avoid wasting time on either facts or values that really have no impact on the decision. Usually this is accomplished via sensitivity analysis.

e. Forces explicit recognition of uncertainty.

DMs often either do not understand or try to duck issues of uncertainty. Analysis forces DM to deal with uncertainty.

f. Understanding of the complete problem.

"the detailed quantitative work in doing the full-scale study probably helped the [DM] to better understand the qualitative implications of the problem, and it was this qualitative understanding that helped him influence the governmental officials." Forces DM "to scrutinize his problem as an organic whole."

Claim 3: Communication

"executives most satisfied with the approach valued it as a vehicle for communicating reasoning amongst decision makers." A number of good outcomes are presumed to follow from this improved communication.

a. Advocacy, defense or decision

"helps DM defend and communicate his decision. . .makes point that decision was not frivolous and that all factors were considered." "provide an effective basis for the justification of a particular recommendation."

3b. Conflict resolution, reconciliation

This claim is made for virtually every approach that aims at application! However, justifications for this claim often are rather murky--conflict resolution seems to happen as a by-product of the improved communication. "useful as a mediating device. . .can quickly focus on those issues on which they have fundamental disagreements. . . .will sharpen the specificity and sophistication of the arguments" "Can spell out explicitly what the values of each participant are, show how much they differ, and in the process can frequently reduce the extent of such differences." "isolate and quantify differences of opinion among experts"

c. Teaching

"The documentation of an analysis can serve as a briefing report for a new DM." Judgment and decision analysis are frequently alluded to as valuable teaching devices for such

diverse disciplines as medicine, airline piloting, and live-stock grading!

d. Explanation and intellectual discussion.

Even if there is no hot conflict, analysis aids an intellectual, non-emotional discussion of critical issues facing society, "can be meaningfully discussed by experts and laymen," "it is easier to discuss the issues with our colleagues."

Claim 4: Policy focus (variable) rather than Specific-Action (object) focus

All approaches, in one way or another, claim that their methods and procedures promote a policy focus, "shift the DM's attention from specific actions to the values these actions serve." The advantages of so doing include the following.

a. Saves time, money, and unhappiness.

"eliminate the need for costly time-consuming case-by-case adversary or negotiating proceedings."

b. Easy to adapt to new information or changing values.

"provides a framework for contingency planning and for the continuing evaluation of new facts that is necessary as the dynamics of a problem unfold," "such policies can easily be changed in response to new circumstances or changing value systems," "provides the opportunity for cumulative knowledge."

c. Easy to pass on to next administrator.

Really same claim as Teaching (3c) and Adapt to New Circumstances (4b).

d. Inform those affected by policy.

"Such policies can be easily, efficiently, and explicitly disseminated," "by publicizing a set of values, a decision making organization can in effect inform those affected by its decisions about its ground rules."

Claim 5: Fact-Value [Probability-Utility] Separation

"distinguishes the DM's preferences for consequences, including his attitudes towards risky situations, from his judgments about uncertainties," "separating judgments of probabilities from those of values permits rational analysis." Some of the advantages that follow from this separation include:

a. To each his own.

"The separation phase permits elected representatives to function exclusively as policy-makers, and scientists to function exclusively as scientists."

b. Easy to integrate information from experts.

"each expert can give testimony in an unambiguous quantitative manner that can be incorporated in the overall analysis," "formal framework makes it easier to incorporate the opinion of another expert."

Claim 6: Synergistic effect, creates new solutions

"aid in generating creative alternatives," "there is always the possibility that these analyses will uncover new insights that result in a different alternative--one that is perceived as better than any of the original alternatives."

[MORE]

METHOD

A. Introduction

This section analyzes the six approaches to judgment and decision making with respect to their methodologies. The distinction between methodology and procedure (the latter being the topic of the next section) is sometimes nebulous. We have tried to include in methodology those issues which pertain to strategic choices for testing or implementing a judgment and decision theory and to reserve tactical choices for the procedures section of this report. For example, the decision to analyze group or individual data is a strategic choice and is discussed in this section, while operational definitions of the key concepts identified in the previous section on theory are included in the procedures section rather than here. Of course, there is inevitably some overlap so please excuse a few redundancies in our quest for systematic coverage of the important dimensions on which the six approaches can or do vary.

The organization of this section is much the same as that of the preceding section: each approach is described in terms of theoretically neutral dimensions and an attempt at integration follows the descriptions for each dimension. The dimensions to be considered are the following.

- B. Idiographic vs. Nomothetic Analysis
- C. Sampling Domains (e.g., subjects, objects, tasks)
- D. Stimulus-Object Decomposition

E. Judgment Decomposition

F. Methods of Partitioning the Decision

Our attempt at integration will involve the same steps employed in the theory section: (a) denotation of similarities and differences (b) denotation of antinomies requiring empirical, logical, or conceptual reconciliation; and (c) steps toward reconciliation whenever possible.

B. Idiographic vs. Nomothetic Analysis

Introduction

In the prototypical judgment and/or decision tasks used in the six approaches, there are m object-judgments from n judges. This yields an $n \times m$ matrix of object-judgments which may be analyzed using idiographic methods, nomothetic methods, or a combination of both. The choice of either an idiographic or nomothetic analysis is extremely important methodologically because it strongly constrains other aspects of study design and limits the generalizability of the results. Before describing each approach's characteristic methodological choice, we briefly describe the four main alternatives.

Purely nomothetic. For the purely nomothetic alternative, n (the number of judges or subjects) is generally large, and m (the number of object-judgments per subject) is small (possibly $m = 1$). The ultimate judgment, result, or test statistic is the average judgment across subjects. If m is greater than 1, then the different object-judgments are treated as replications so that there are m average judgments. That is, the $n \times m$ matrix is collapsed into a single row of object-judgments and differences between judges are treated as random error.

Pure idiographic. The purely idiographic alternative is obviously the opposite of the purely nomothetic one; n is generally small (possibly $n = 1$) and m large. The quantitative result is the average judgment for each person or a summary measure for each person's set of judgments indicating whether or not the judgments fit a particular model or not or separate model parameter estimates for each individual.

These measures may be statistical deviations from a model (e.g., R^2 , F for an interaction term) or a nominal code (e.g., 1 = "show bias," 0 = "no bias"). If n is greater than 1, then the different subjects are treated as separate replications so that there are n average judgments, summary measures, or parameter estimates--one for each subject. That is, the $n \times m$ matrix is collapsed into a single column and differences between object-judgments (or deviations of those judgments from a model) are treated as error.

Idiographic/nomothetic combination. For this combination of approaches, both n and m tend to be reasonably large (although that is not essential). The critical point is that the idiographic analysis is done first and the results are then averaged across subjects in a nomothetic analysis. Model fitting, parameter estimation, statistical tests, etc. are carried out separately for each judge or subject and it is those values, rather than the raw judgments, that are averaged or counted (e.g., "X% violated additivity") across subjects. In this case, the $n \times m$ matrix is collapsed into one column and then into a single indicator. The main distinction between the purely idiographic and the idiographic/nomothetic approaches is that in the former results for different

individuals are viewed as complete, independent replications (in the same sense that one whole experiment replicates another) whereas in the latter results for different individuals are averaged together in order to define a "composite" subject.

Nomothetic/idiographic combination. The name of the alternative strains the meaning of "idiographic" but because the sequence of steps in this alternative is exactly opposite to that of the preceding "idiographic/nomothetic combination" the label seems appropriate. Both n and m also tend to be reasonably large in the nomothetic/idiographic combination. However, the order of the nomothetic and idiographic steps is reversed. The first step is to average judgments across subjects and then to submit those average judgments to further analysis (such as fit to a linear model). Model fitting, parameter estimation, statistical tests, etc. are done only for the aggregate judgments. In this case, the $n \times m$ matrix is collapsed into one row and then into a single indicator.

We now turn to a classification of the characteristic methods of each of the six approaches into one or more of these prototypical methodological alternatives.

Decision Theory

The primary methods of DT are easily classified into the purely idiographic category. In most cases there is only one judge or decision maker so that there is no opportunity for nomothetic analysis. This methodological choice obviously derives from theory. Because probabilities and utilities (the important parameters) are presumed to be personal and subjective there is no expectation that differences between people are due only to random error as is required by the nomothetic approach. Rather the expectation is that people will have different parameters and thus their judgments will

require individual analysis. Nomothetic analysis simply makes no sense in the context of the theory.

The one minor exception to the pure idiographic method is Keeney and Raiffa's treatment of group decision making. The "supra Decision Maker" model (see Chapter 10 of Keeney and Raiffa, 1976) for aggregating individual utilities into a joint decision is an example of the idiographic/nomothetic methodological approach. Individual utility functions are determined first and then combined into a single utility using the Decision Maker's weights for each individual. However, this method does not seem to have been used in any applications.

Behavioral Decision Theory

A variety of methodological approaches have appeared in the work of BDT. The earliest studies within this tradition relied mostly on the purely idiographic method (e.g., Mosteller & Nogee, 1951; Edwards, 1953, 1954; Davidson, Suppes, & Siegel, 1957). In these studies there are generally separate results, figures, analyses, etc. for each subject (or at least for a selected sample of subjects).

The "conservatism" studies of probabilistic information processing often abandoned the pure idiographic approach for the nomothetic/idiographic combination. For example, in the three experiments reported in Phillips & Edwards (1966) the probability estimates from several (5 to 12) subjects are averaged and then those average estimates are compared to the probabilities predicted

by Bayes' Rule. Parameter estimates follow the aggregation across subjects (e.g., in the equation $\Omega_1 = L^c \Omega_0$ the parameter c is estimated only for the group data).

Finally, the most recent work on MAUT (the SMART method described in Gardiner & Edwards, 1975) has returned to the use of the idiographic approach (see their Figure 5) although some nomothetic results are also reported (see their Figure 6).

Psychological Decision Theory

The work within PDT also exhibits a variety of methodological approaches; it is not unusual for both idiographic and nomothetic methods to be used in the same study. To complicate matters further it is often difficult to categorize some PDT methods as being either idiographic or nomothetic.

Recent experiments on "prospect theory" (Kahneman & Tversky, 1977) provide good examples of PDT's use of nomothetic methods. The following is a typical result from that paper in which subject behavior is compared with SEU predictions.

Percent selecting gamble A over gamble B: 67%

Percent selecting gamble A' over gamble B': 22%

The gambles are constructed so that in order to be consistent with SEU axioms, a choice of gamble A implies choice of gamble A' and a choice of gamble B implies choice of gamble B' (as well as both converses). Subject behavior is apparently not in agreement with such implications because a majority choose A in the first pair and B' in the second pair. This is the nomothetic approach because

there is never a count of exactly how many individuals switch from A to B' or B to A' (in the present example the proportion of individuals violating the SEU predictions could range from a low of 45% to a high of 89%).

On the other hand, Tversky and Kahneman (1977) in their paper on "Causal Schemata" use an idiographic/nomothetic combination using the same two-choice paradigm as employed in the prospect theory studies. Instead of reporting choice percentages for each pair, they report "A large majority of our subjects (56 or 68) stated that $P(A/B) > P(A/\bar{B})$ and that $P(B/A) < P(B/\bar{A})$, contrary to the laws of probability" (p. 10). This is idiographic analysis, for it is clear that most individuals do indeed make simultaneous, contradictory choices.

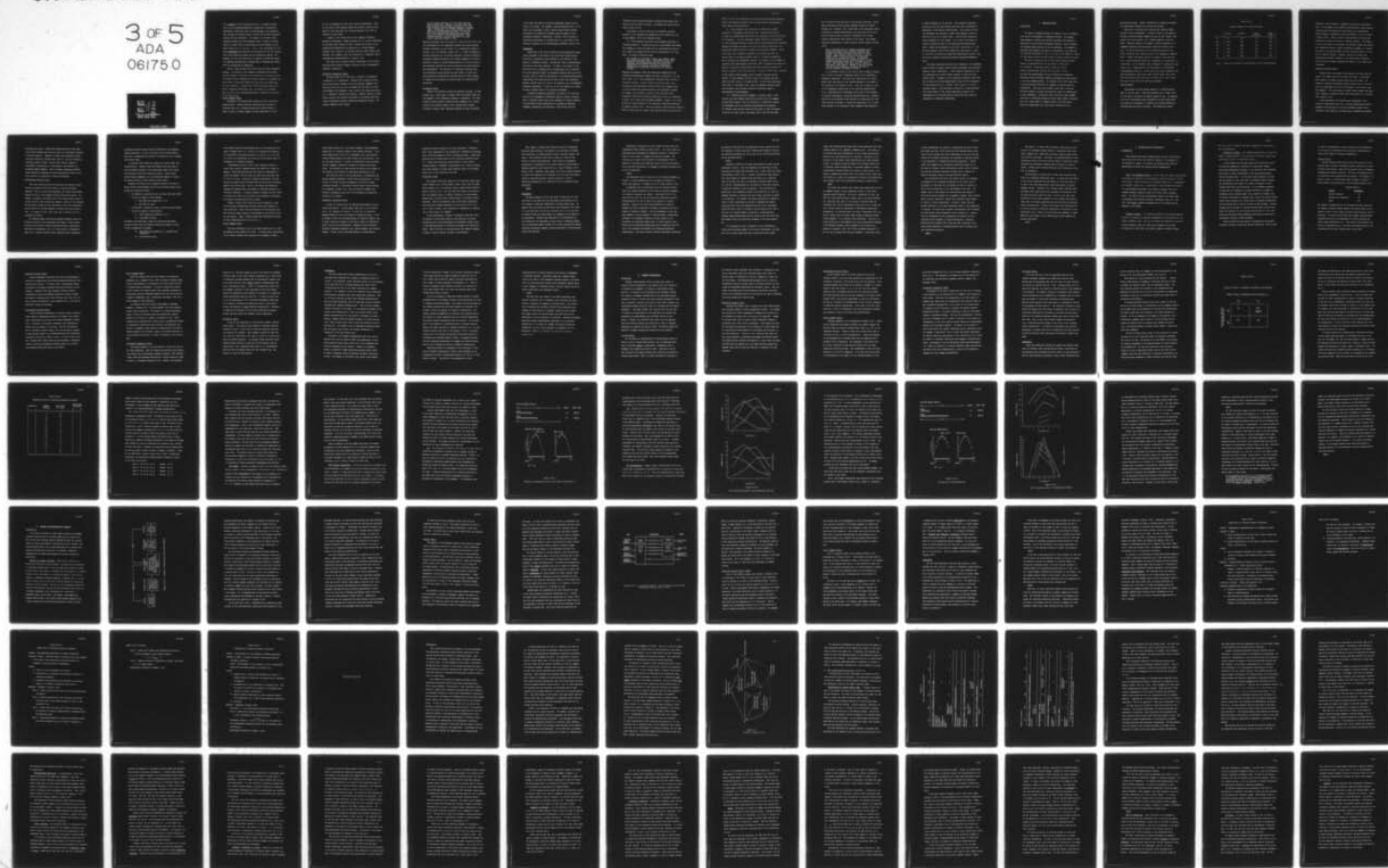
The methods of the heuristic studies (e.g., Kahneman & Tversky, 1972, and Tversky & Kahneman, 1973) are not so easily classifiable. Classification is difficult because although the typical study involves large subject samples with only one subject-judgment or choice per subject--the usual large n, small m pattern characteristic of nomothetic methods--the data are often (but not always) analyzed using an idiographic/nomothetic combination. That is, on the basis of a few responses (and often only a single response) each individual is categorized as either exhibiting or not exhibiting the given error or bias and then the numbers in each category are counted. For example, in Tversky and Kahneman (1973) 105 of 152 subjects evidenced an availability bias by incorrectly believing that more words had a given letter in the first than in the third position

AD-A061 750

COLORADO UNIV BOULDER CENTER FOR RESEARCH ON JUDGME--ETC F/G 5/10
THE COLORADO REPORT ON THE INTEGRATION OF APPROACHES TO JUDGMEN--ETC(U)
OCT 78 K R HAMMOND, G H MCCLELLAND N00014-77-C-0336
CRJP-213 NL

UNCLASSIFIED

3 OF 5
ADA
06175 0



for a majority of five letters (K,L,N,R,V). It should be noted that in such idiographic/nomothetic combinations in heuristics experiments, statistical tests are applied only to the nomothetic part although in principle there is nothing to preclude statistical analyses for each individual. For example, in the above letter frequency problem, the null hypothesis (that a subject is equally likely to choose first or third position as more frequent for any given letter) can be rejected (at $p < .05$, one-tailed) for only those individuals who choose the first position for all five letters. Still other heuristics studies use pure nomothetic methods. For example, the fact that sample size is ignored in estimating subjective sampling distributions is demonstrated in Kahneman and Tversky (1972) by between-group comparisons.

Finally, there are many PDT studies which use pure idiographic methods. In contrast to the nomothetic experiments which demonstrated systematic violations of SEU axioms and the consequent need for Prospect Theory, Kahneman and Tversky estimate the parameters for their theory (uncertainty weights and value functions) separately for each individual. Tversky (1967, 1969, 1972a) also employs the purely idiographic methodology (e.g., see Figure 10 in Tversky, 1967) as well as the idiographic/nomothetic combination (see Table 3 in the same source).

Social Judgment Theory

Proponents of SJT specifically advocate the use of the idiographic method. Because different individuals may use cues in different ways (i.e., different weights and function forms) it makes no sense to average judgments across individuals or to fit

one set of parameters for data from a group of individuals. Thus, in the typical study separate weights and function forms are estimated for each individual and a separate measure of fit (R^2) is computed for each person.

A number of SJT studies also use the combined idiographic nomothetic methodology. Summary measures (R^2 , G) for each individual are averaged within groups in order to assess the effects of such between-group manipulations as cognitive vs. outcome feedback (Adelman, 1977), linear vs. curvilinear function forms in MCPL tasks (Brehmer, 1974; Hammond, 1971) and correlated vs. orthogonal cues in interpersonal learning (Mumpower & Hammond, 1974).

The recent use of cluster analysis (Rohrbaugh, 1977) to group individuals with similar judgment policies is also an idiographic/nomothetic combination.

Information Integration Theory

The work within IIT is difficult to classify as idiographic or nomothetic or a combination. Anderson (1974) explicitly advocates within-subject designs, but in such designs judgments from different individuals are still averaged when the model is tested and parameters are estimated. Thus, in most IIT studies only group (i.e., nomothetic) analyses are reported. However, reports of those studies usually also include a disclaimer to the effect that individual (idiographic) analyses confirm the nomothetic results. For example, Shanteau (1975) states:

All the results reported up to this point have been based on group data. However, it is well known that group analyses can mask model discrepancies for individual subjects (see, for example, Shanteau & Anderson, 1969). Accordingly, each of these studies was also analyzed at the single-subject level. For the most part, these analyses have substantiated the group results in support of the multiplying model. Where discrepancies have occurred, they tend to be small with no consistent trends (p. 118).

When group analyses might mask important individual differences then IIT investigators do use idiographic analyses and report separate parameter estimates for each subject or a representative sample of subjects (e.g., see Shanteau & Nagy, 1976). It seems reasonable to conclude therefore that IIT prefers nomothetic (actually nomothetic/idiographic combination because the averaged judgments are further analyzed in terms of a model) methodologies but is quite willing to use idiographic methods when the data indicate that group analyses are inappropriate. The preference for nomothetic methods stems from convenience (group analyses are much easier to report) and from statistical considerations (grouped data have larger degrees of freedom and much greater statistical power for detecting model deviations).

Attribution Theory

Almost all attribution studies use nomothetic methods. In some studies (e.g., Jones, Davis, & Gergen, 1961) the subjects make only one judgment so that idiographic analysis is impossible. For those studies in which subjects do make multiple judgments (e.g., Weiner & Kukla, 1970; Frieze & Weiner, 1971; and many other studies reviewed in Frieze, 1976), nomothetic methods are used exclusively.

If one looks very hard an occasional idiographic analysis can be found in AT studies. For example, Jones and Goethals (1971, p. 36) report the percentages of their subjects using various distinct strategies in an impression formation task--an example of the idiographic/nomothetic combination. However, even in studies of "individual differences" in attributions (e.g., see Frieze, 1976) nomothetic analysis is the overwhelmingly predominant choice of AT.

Integration

Unlike many of the other theoretical and methodological dimensions considered in this report, the distinction between Group I and Group II approaches is not reflected in the choices of idiographic or nomothetic methods. Instead both types of methodologies are found in each basic approach. In terms of this issue, we can match each Group I approach with a corresponding Group II approach. DT and SJT generally depend on idiographic analysis; BDT and IIT are the most likely of these six approaches to use nomothetic/idiographic combinations; PDT often uses and AT almost always uses nomothetic methods; and BDT, PDT, SJT, and IIT all sometimes use the idiographic/nomothetic combination. So the use of all four methods can already be found on either side of the conceptual watershed.

Is further integration necessary? We think so, for often the choice of an idiographic or nomothetic method is based on convenience or tradition rather than on any conceptual or logical analysis of which method is more appropriate for studying or measuring judgments, preferences, and decisions. Only SJT and IIT have

thought the issue important enough to advocate which method type should be used in what situations. (One might also argue that the issue is moot for DT.)

This report is not the place for the complete conceptual analysis of the strengths and weaknesses of each method type, but we can suggest a few of the important issues.

(1.) Group analyses can yield very misleading indicators of individual behavior. Grouped data may fit a given model even though the judgments of every individual do not, and grouped data may not fit a model even though the data from all individuals do fit that model. As Tversky (1972a) has noted:

most studies [of preference, choice, and judgment] report and analyze only group data. Unfortunately, group data usually do not permit adequate testing of theories of individual choice behavior because, in general, the compatibility of such data with theory is neither a necessary nor a sufficient condition for its validity (p. 291).

Shanteau and Anderson (1969) have empirically demonstrated such problems and hypothetical examples are easy to construct (e.g., see Luce, 1959, p. 8). This consideration obviously favors the use of idiographic methods. Decision tasks involving different value structures across individuals are more likely than judgment tasks to reveal differences between nomothetic and idiographic methods. For example, most instances in which IIT has reported purely idiographic results have involved preferential choice (of sandwiches, of dates, etc.) rather than knowledge judgments. However, it is quite possible for group analyses to hide important individual differences in judgment tasks as well. Thus, any approach using nomothetic analysis should be aware that the "group" model derived from such analysis may

apply to none of the individuals in the group when considered separately; ideally any nomothetic analysis would include specific methodological checks against that possibility.

(2.) The strength of any conclusions is affected by whether nomothetic or idiographic methods are used. Unfortunately there is at least one consideration which favors the nomothetic and one that favors the idiographic. With an idiographic method, analysis is concentrated on a few individuals and one is never sure how representative the resulting parameter estimates and model fits are of the general population; thus populational generality may be stronger with nomothetic methods. On the other hand, with a nomothetic analysis we are never sure that the results apply to any one individual; that is, situational generality may be stronger with idiographic methods. For example, in the example of nomothetic data from Kahneman and Tversky (1977) discussed above while the majority of subjects selected A from the first pair and a majority of subjects choose B' from the second pair, it does not follow that a majority of the subjects selected both A and B' thereby violating the SEU axioms. In this instance, as few as 45% of the subjects may have made this joint choice, or as many as 89%. If the actual percentage were closer to the higher value, then an idiographic analysis would have produced a much stronger conclusion relatively immune to methodological second-guessing.

(3.) Presenting multiple judgment or decision tasks to each subject may create demand characteristics which alter the judgment process being studied. This is particularly a problem for certain AT experiments which use elaborate manipulations and scenarios. For example, most subjects would have been able to infer the manipulation in the Jones, Davis, and Gergen (1961) study had they been

able to listen to more than one of the recorded interviews. While some investigators using multiple judgments dismiss the demand problem (e.g., Anderson, 1974, claims many of his experiments which controlled for demand characteristics and those that did not produced the same pattern of results), remarkably little formal research attention has been devoted to this issue. The dilemma facing any investigator is aptly stated by Ostrom, Werner, and Saks (1977).

One of the difficulties in drawing conclusions in the present experiment is that parameter estimates were obtained on a group wide basis rather than for individuals. This resulted from the decision to counter-balance stimuli across set sizes on a group basis rather than on an individual basis. If parameter estimates are desired on an individual basis, it would require having each subject see all stimuli in all set sizes. Each stimulus item, then, would be seen more than once by each subject. Whether or not this design alteration would change the processes involved in juristic judgments is unknown at present.

It is probably obvious by now that we favor, primarily because of (1.) above, the purely idiographic method or the idiographic/nomothetic combination. But within the combination, the investigators still must decide on the relative emphasis to be placed on the idiographic and nomothetic parts. Thus, the strength and type of the idiographic analysis may be very weak and nonstatistical (e.g., classification of individuals on the basis of one response in some PDT experiments), strong and nonstatistical (e.g., examination of all an individual's many choices for consistency with each other and with DT axioms), or strong and statistical (e.g., statistical analysis of an individual's many judgments using regression

or ANOVA techniques by SJT and IIT). The subsequent nomothetic statistical test may then involve relatively weak, nonparametric techniques or more powerful ANOVA methods. Such choices for both the idiographic and nomothetic stages raise immediate questions (which we do not try to answer here). Should formal statistical tests be employed for both stages? If so, what is the effect of "double" statistical tests on the experiment-wide significance level? Should the idiographic test be very liberal (say, $p < .30$) and the nomothetic test be more conservative (the usual $p < .05$), or vice-versa, or should the respective significance be equal? And how would a Bayesian statistician analyze idiographic/nomothetic data?

The issues discussed above are just a sampling of the considerations involved in making a choice between idiographic and nomothetic methods or their combinations. A more complete integration of the six approaches will require further analysis of these and other related issues. Agreement on the idiographic, nomothetic choice may well be the most critical methodological issue for the integration of these approaches (e.g., Brunswik, 1952, raised it to an all-important issue). While awaiting a resolution, it seems important that practitioners of the various approaches be aware of the consequences of the choices they are making when they select an idiographic or nomothetic methodology.

C. Sampling Domains

Introduction

The choice of sampling domains is related to but not redundant with choice of idiographic or nomothetic methods. For example, large samples of subjects are usually associated with nomothetic methods but do not preclude idiographic analysis. Here we examine the sampling domain choices of the six approaches. No judgment or decision study presents one object to one subject in one task, but the different approaches do vary in the size of the samples or number of replications allotted to each domain of subject, object and task.

The size of sample for any of these three domains obviously affects the generalizability of the results. Often there is a tradeoff: using many objects or tasks in a study inhibits sampling many subjects and vice-versa. We can see that this methodological choice is important by examining instances in which different sampling strategies have been applied to the same judgment problem. For example, Cliff (1959) found that a multiplying model described the judgments of adverb-adjective combinations. Cliff used large samples (about 200) of subjects and small samples of adverbs and adjectives (about 40 combinations in each experiment). McClelland (1974) studied the same judgment task using a contrasting sampling strategy--small sample of subjects (21) and a large sample of judgment objects (over 200 adverb-adjective combinations) and found serious violations of the

multiplicative model. Clearly, differences in sampling strategies can dramatically influence the conclusions obtained.

Table III-C-1 indicates the characteristic sampling domain choices of the six approaches. "Subjects" refers to the number of subjects or decision makers sampled (we use the word "sample" very loosely and do not restrict it to "random" samples) in any one study. "Objects" refers to the number of objects presented to or judgments required of any one subject within the context of one task. We have divided tasks into two domains--substantive and formal variations (this distinction was first suggested by Hammond, 1966). "Formal" variations pertain to changes in task structure (e.g., intercorrelations among cues, redundancy or inconsistency of information; type of response scale) independent of substantive task content (e.g., judging people, lifting weights, or establishing land-management policies). Unlike object and subject sampling, task sampling generally takes place across studies rather than within. Below we briefly justify each entry in Table III-C-1 and indicate the nature of the particular sampling strategies.

Decision Theory

The methods of DT are usually applied to a single decision maker in any one study. There are exceptions (e.g., Roche, 1971) but the norm is certainly for subject samples of one. The emphasis in DT is on collecting as much information from that single decision maker as is necessary to construct the required probability distributions and utility functions. This method may require

Table III-C-1

Sampling Domains of the Six Approaches

| | Subjects | Objects | Tasks Substantive | Tasks Formal |
|-----|----------|---------|----------------------|-----------------|
| DT | No | Yes! | Yes | Some |
| BDT | Some | Yes | Yes | Some |
| PDT | Yes | Some | Yes! | Some |
| SJT | Some | Yes | Yes | Yes! |
| ITT | Some | Yes | Yes! | Yes |
| AT | Yes! | Some | Yes | Some |

Note: ! indicates an unusually strong emphasis on that sampling domain.

hundreds or even thousands of judgments and decisions (see Keeney, 1977, for an example of an extensive interview with a single person). Across studies, DT methods have been applied to an impressively diverse sample of substantive tasks (see Chapters 7 and 8 of Keeney & Raiffa, 1976). Finally, there is some variation of formal task properties in terms of the type of response required. For example, to test the validity of the axioms, the decision analyst may present pair comparison choices, request indifference judgments, or explain the axiom to the decision maker and ask the decision maker to evaluate its validity directly. However, selection of formal task characteristics seems to depend on the situation and convenience and there are no studies evaluating the effect of such formal task variations.

Behavioral Decision Theory

Studies within the context of BDT generally use more than one subject, but sample sizes are typically not large. On the other hand, samples of stimulus objects tend to be quite large. For example, in Phillips and Edwards (1966) there are from 5 to 12 subjects per experimental group and in Experiment I each subject made 480 judgments. This same pattern of small subject samples and large object samples is repeated in many other BDT studies (e.g., Fryback, 1974; Fischer, 1977).

Across experiments, BDT studies have investigated a wide variety of substantive tasks (e.g., military intelligence, medical diagnosis) although some of the early experiments showed little substantive task sampling by concentrating on gambling and bookbag-

and-poker-chip tasks. Formal task characteristics are also sometimes varied although this does not seem to be a particular interest of the BDT approach. Examples of formal task variations include different probability response modes (odds vs. log odds; bounded vs. unbounded response scales), scoring rules (linear, quadratic, logarithmic), hypothetical vs. real gambles, and wholistic vs. decomposed SMART judgments. Work in progress comparing different formal methods for assessing the same multiattribute utility structures can also be classified as formal task sampling.

Psychological Decision Theory

The recent studies within PDT (heuristics and prospect theory) reverse the pattern of BDT with respect to subject and object sampling. That is, typical PDT experiments use rather large subject samples with only a few judgments collected from each subject. For example, in Kahneman and Tversky (1977) each subject made at most 12 choices and each choice problem was judged by from 66 to 115 subjects. The heuristics studies show the same pattern, but a few nonheuristic studies switch the relative emphasis on subject and object replication. For example, Tversky (1967) used only 11 subjects with 225 responses per subject.

The sampling domain receiving the greatest emphasis within the heuristics studies is the substantive task domain. These studies illustrate the heuristic biases (representativeness, availability, anchoring with adjustment, etc.) in a wide variety of substantive tasks (e.g., judging occupation types, guessing letter frequencies,

estimating binomial sequence relative likelihoods, and answering almanac questions). In fact, the point of such studies seems to have been to demonstrate the ubiquity of heuristics in all judgment and decision tasks.

In contrast, PDT studies have usually not varied formal task characteristics. However, those few studies which have done so provide dramatic evidence of the considerable impact that formal task variations may have on empirical results. Because those studies figure prominently in our integrative comments to follow, we consider several of these studies in some detail.

Kahneman and Tversky (1977) presented subjects with pairs of paired choices between gambles such as the following problem illustrating the "isolation" effect.

1. In addition to whatever you own, you have been given \$1000.

You are now asked to choose between

A: (win \$1000 with probability .5) or

B: (win \$500 for sure).

2. In addition to whatever you own, you have been given \$2000.

You are now asked to choose between

A': (lose \$1000 with probability .5) or

B': (lose \$500 for sure).

According to the SEU axioms, these two choices are equivalent because in either case the effective choice when viewed in terms of final consequences is between

A'': (win \$2000 with probability .5; otherwise win \$1000) and

B'': (win \$1500 for sure).

This formal variation--the appearance but not the substance of the gambles changes from A to A' and B to B'--reverses the preference of the majority; the majority prefer B to A and A' to B' (but see the caution in interpreting this result in the discussion above of idiographic vs. nomothetic methods).

Lichtenstein and Slovic (1971) used a bidding procedure to obtain their subjects' maximum buying price (MBP) for a number of gambles. These same subjects were also given the opportunity to choose that gamble (from the same set) which they would most like to "own" or to play. Contrary to axiomatic predictions, for many subjects the gamble with the highest MBP was not the most preferred gamble in the choice task! That is, the formal task variation--changing the response mode from bidding to choosing--resulted in a different preference. A detailed attempt by Grether and Plott (1978) to explain this result as an artifact only succeeded in confirming the original Lichtenstein and Slovic result.

Finally, Tversky (1967) found that utility assessed in riskless situations systematically differed from utility measured in risky situations. Risky utilities exceeded the corresponding riskless utilities (again contrary to SEU axioms) at all points for all but one subject. Again, a formal change which theoretically should not have changed behavior did in fact do so.

Social Judgment Theory

The clear preference of SJT is for small sample sizes (1 to 10) and moderate object samples (20 to 50). In recent policy applications of SJT "subject samples" have consisted of all members of small,

identifiable groups (e.g., city council members, labor-management negotiators in a specific dispute, state cabinet officers). A few survey applications of SJT (e.g., Rohrbaugh, 1977) have involved several hundred subjects but these studies are the exception. SJT has also been applied to a variety of substantive topics including clinical judgment, perception, probability learning, and public policy. The diversity of topics within the public policy domain is very similar to the variety of substantive applications of DT.

More than any other of the six approaches, SJT emphasizes and employs extensive sampling of formal task characteristics. In fact in many of the MCPL studies the emphasis is so much on formal properties (linear vs. curvilinear function forms, intercorrelated vs. orthogonal cue sets, etc.) that the specific judgment task includes a substantive topic only for the purpose of holding the interest of the subject.

Information Integration Theory

In most IIT studies there are moderate-sized samples of both objects and subjects. Typical sample sizes are 10 to 15 subjects and 9 to 25 judgments per subject. As with PDT, the important sampling domain for IIT is the domain of substantive tasks. IIT in an attempt to show the wide applicability of cognitive algebra, has, it seems, applied its methods to the widest variety of substantive tasks of any of the six approaches. Examples include perceptual illusions, impression formation, jury decision making, and livestock judging. In fact, IIT has even been applied to characteristic

substantive tasks of several of the other approaches: Shanteau's (1970, 1972) experiments on the probabilistic information processing tasks of BDT and Anderson's (1974) attribution studies. Formal task properties have received less but still important attention within IIT. Formal task variations include studies of order effects (primacy-recency), redundancy, inconsistency, etc. Notice, however, that these analyses of formal task properties refer to the stimuli proper, not to the structure of the task.

Attribution Theory

Most studies within the tradition of AT have used rather large subject samples (e.g., 258 in Weiner & Kukla, 1970), and very small object samples (e.g., 118 subjects judged a single object-person in Nisbett & Wilson, 1977). Not only does each subject often receive only one judgment object or stimulus condition, but even between groups there is usually only a limited number of different stimuli presented. There are exceptions; some of the Weiner experiments have used object samples comparable to those used by SJT or IIT (e.g., Frieze & Weiner, 1971, presented 54 combinations of information to each subject for judgment).

AT has investigated a variety of substantive tasks but often the purpose of this variety seems to have been more to provide interesting (for social psychologists) experimental contexts than to show generalizability of attribution processes over substantive areas. Some of the work on self-attribution has purposely sampled a number of topics (obesity, insomnia, psychotherapy).

With respect to formal task variations, Kelley's covariation principle leads directly to an interest in the effects of information consistency (a formal property) across occasions, objects, and actors. Thus, McArthur (1972) may be viewed as a study of the effects of formal task variations. Order effects in sequential presentation of information is one formal property that has been an important empirical topic within AT (e.g., Jones & Goethals, 1972: Frieze, 1976). Otherwise, there seems to be little attention devoted to formal task sampling in AT (although we are not sure that manipulations of perceived decision freedom, success vs. failure, behavioral desirability, etc. should not also be considered formal variations).

[MORE]

Integration

The high frequency of "yes" and "some" in Table III-C-1 indicates that in the sense that all approaches think sampling in all four domains is important integration is already taking place. We suspect that the picture is even better than indicated by the table. Our first table had very few yes's and many more no's but the more we looked the more we found evidence for sampling in all domains by all approaches. Probably more indicative of the differences (if any) between the six approaches in their choice of sampling domains are the exclamation marks in Table III-C-1 which indicate the domains receiving the greatest emphasis (either empirically or theoretically) within each approach.

Particularly striking is the full column of yes'es under the "Substantive Tasks" heading in Table III-C-1. As a body, these six approaches have investigated an impressive variety of substantive topics in the context of judgment and decision problems. The diversity is so great--from economics to psychotherapy--that it seems as if virtually every important topic in the social sciences can or has been interpreted as a judgment or decision problem by one of the six approaches.

[MORE]

The differences that do exist are (a) the relative emphasis on subject vs. object sampling and (b) the perceived necessity of formal task sampling. We comment on each of these issues in turn.

We emphasize that all approaches use both multiple subjects (with the possible exception of DT) as well as multiple judgments (or choices or whatever) per subject. From Table III-C-1 it is clear that DT, BDT, SJT, and IIT place more importance on object sampling whereas PDT and AT devote more attention to subject sampling. (Again we note that these differences in relative emphasis are mostly independent of whether an approach uses idiographic or nomothetic methods. The only dependence is that very small samples of either subjects or objects inhibit, respectively, nomothetic or idiographic analysis; yet many studies in all approaches use sufficient sample sizes for either analysis type. Within the four approaches emphasizing object sampling we find the use of both nomothetic/idiographic and idiographic/nomothetic combinations.) This issue certainly deserves additional theoretical

and empirical work from the six approaches because varying the relative emphasis on subject and object sampling can change the results and conclusions obtained (as in the case of the Cliff and McClelland studies discussed above).

[MORE]

Systematic variations of formal task properties across the six approaches is spotty. Only SJT emphasizes the importance of such variations and we gave most approaches only a "some" under the formal task heading of Table III-C-1. Further, although the same substantive topics appear in studies across approaches (e.g., gambling tasks in DT, BDT, PDT, and IIT; impression formation tasks in SJT, IIT, and AT), experimentation on any specific formal task variation is generally limited to only one approach. Such specificity is partly due to the theoretical bases of each approach for often the language necessary to describe any given formal variation is unique to one approach. For example, the formal equivalence of gambles in terms of final consequences as studied by Kahneman and Tversky (discussed above) only has meaning within the Group I approaches which have their origins in axiomatic utility theories (DT, BDT, & PDT) and, for another example, variations of intercorrelation patterns among informational cues as studied by SJT are really only describable within the theoretical framework of SJT (and possibly AT).

Is it important to seek a resolution of the differences in formal task variations sampled by the various approaches? We think so if for no other reason than most studies which have studied

formal task properties have found them to have significant and often dramatic impacts (e.g., Mumpower & Hammond, 1974). (Of course, it is possible that many investigations of such formal manipulations which did not find effects were not reported.) Reactions to the existence of such potent formal manipulations have been extremely varied. For example, Coombs and Bowen (1971) dismiss effects due to different formal presentations of equivalent gambles as rather uninteresting "display effects" not central to basic theory. Conversely, Kahneman and Tversky (1977) and Tversky (1972a, 1972b) use such effects as a basis for discarding old theories and building new ones (prospect theory and elimination-by-aspects, respectively).

[MORE]

One formal task property that varies across approaches--the use of judgment ratings vs. pair comparison choices or rank order--deserves special attention. The Group I approaches (DT, BDT, and PDT) with their decision theory origins tend to use choice methods for collecting data; that is, the subject or decision maker is confronted with various alternatives (often only two at a time) and must indicate either his preference or indifference among these alternatives. On the other hand, the Group II approaches (SJT, IIT, and AT) usually have the subject make ratings on numerical scales (often presumed to have interval scale properties). There are of course exceptions to that general rule such as probability estimates in BDT Bayesian studies, choice problems in IIT studies (e.g., Anderson & Alexander, 1971), and "direct preference measures" in DT, but let us proceed with the basic argument. Given such a split

in basic methodology, the results of Lichtenstein and Slovic (1971) and Grether and Plott (1978) which show that maximum buying prices (akin to worth ratings) and choice preferences do not give the same results are extremely disturbing, and represent an important barrier to the integration of judgment and decision approaches. Tversky (1969) also suggests that considering alternatives singly (as in the case for ratings) versus alternatives in pairs (as for choices) may activate different evaluation processes (e.g., additive vs. additive difference models) producing different results.

The Lichtenstein and Slovic and Tversky results are extremely troublesome for any attempted integration of the Group I and II approaches for they mean that the methods (as well as results) of one of the two groups may not be transferable to the other. For example, must we ask whether SJT rating methods when applied to choice and decision problems (as in Hammond & Adelman, 1976) actually select the most "desirable" outcome (with desirable being defined in terms of a choice methodology)? We need to answer that and similar questions for the other approaches before a true integration is possible. We should note, however, that Shanteau (1972) obtained the same results using preferential choice ratings (i.e., subject rates his preference for one alternative over another) as with subjective-worth judgments of single alternatives. Similarly, McClelland and Rohrbaugh (1977) found the same violations of the Pareto axiom regardless of whether subjects rated or ranked potential arbitration decisions.

[MORE]

The potency of formal task variations, particularly as compared to the seemingly minimal effects of substantive task variations, suggests that formal task properties require greater attention from all approaches. Conversely, the generalizability of both methods and results over such a diverse range of substantive topics indicates that substantive sampling need no longer be a goal in itself. Further work on formal variations would certainly aid the work of integration.

The prevalence of effects due to formal task variations also bears on an important theoretical issue. It argues that the models of the various approaches are more models of the formal tasks (or at best of the task-judge system) rather than models of people's judgment process. Brunswik (1952), Brehmer (1969), Edwards (1971) and Dawes (1975) among others have discussed the possibility that our models primarily model the task. If such is the case, then differences in the formal properties of the judgment and decision tasks used by different approaches are much more than just methodological idiosyncracies. Rather, these formal task differences may be the key barriers to effective integration; a resolution of such differences (e.g., a theory of both choices and judgment) is fundamental for any attempt to find common ground across judgment and decision approaches.

[MORE]

D. Stimulus-Object Decomposition

Introduction

Each approach must make a methodological (as well as procedural) choice about the form of the stimuli or objects to be presented to the subjects or judges. We identify the following three distinct methodological choices for stimulus-object decomposition.

Whole, non-decomposed objects. In this case, the subject sees complete real objects as opposed to the hypothetical, abstract stimuli to be described below. These objects may be relatively simple (e.g., a photograph) or extremely complex (e.g., a description of the procedures and results of a psychological experiment as in Nisbett & Borgida, 1975, or a complete scenario recorded on tape as in Jones, Davis, & Gergen, 1961). For there to be any interpretable results, the objects must be decomposed into relevant variables, attributes, cues, etc. However, these prior conceptual decompositions of the stimuli are not apparent to the subject.

Schematic stimuli. The obvious alternative to presenting complete objects is to display a conceptual decomposition of these objects to the judge in schematic form. That is, the stimuli are described in terms of the key variables, cues, etc. that are of interest in the study. So instead of a real object, the subject judges (or selects between

two, etc.) only a "profile" that may correspond to a real object or may be hypothetical.

Variables as stimuli. The ultimate decomposition is to ask the judge to respond directly to the variables of interest without ever presenting any real or hypothetical objects.

A brief example may help make these three distinctions clear. To study the judgments or decisions of a high school student evaluating undergraduate colleges, we can construct three different basic kinds of tasks corresponding to the three methodological choices described above. For the whole object approach, we could ask the student to rate, rank order, or whatever various colleges and universities after visits to a number of campuses or after viewing films about each institution. In the schematic approach, we would ask for responses to colleges described on a number of variables such as student/teacher ratio, number of students, size of library, and quality of football team. Except for profile values across those variables the colleges would be otherwise unidentified and the profiles may or may not be based on real colleges. Finally, we might ask the student to indicate directly which of those variables are important for the decision and which are not or to rank order the variables without presenting descriptions of any real or hypothetical college on those variables.

While we think the three types of decomposition are mutually exclusive, we make no claim that they are exhaustive. They do seem

to cover the methodological variety among the six approaches and we now turn to a description of each approach's methodological choices with respect to stimulus decomposition.

Decision Theory

Decision analysts almost always pose their questions to the decision maker in terms of schematic objects, especially in multi-attribute applications. There are two distinct types of schematic decompositions employed within DT. The first is decomposition of an alternative into probabilities and outcomes. For example, in a medical context, a decision maker might be asked to consider a given surgical option described in the following schematic form:

Surgical Option X

| <u>Outcome</u> | <u>Probability</u> |
|--------------------------|--------------------|
| Complete recovery | .50 |
| Recovery with disability | .30 |
| Death | .20 |

The second is decomposition of all outcomes into many attributes. For example, different sewage disposal alternatives might be described in terms of costs, water pollution, land pollution, and air pollution. It should be noted that such schematic multiattribute decompositions used in the assessment of utility functions are almost always hypothetical (i.e., they have no real referent) and often very implausible (e.g., a profile with maximum values on all attributes but one with a minimum value on that one).

Behavioral Decision Theory

Various experiments within BDT have used all methodological choices for decomposition; the following examples demonstrate this methodological variety. In Fryback (1974), undecomposed stimuli consisting of intravenous pyelogram films were presented to radiologists. Gardiner (1974) used schematic stimulus profiles (hypothetical coastal developments described in terms of number and density of dwelling units, feet from high tide line, etc.) as well as direct presentation of the variables only, to be used for importance ratings and rankings.

Psychological Decision Theory

In the heuristics experiments, the subject usually receives a rather complex, undecomposed stimulus--for example, the "personality" descriptions in the base rate studies (Kahneman & Tversky, 1973). Certainly, the important variables which differentiate stimuli are not apparent to the judge. (But see Lichtenstein, Slovic, Fischhoff, Layman & Combs, 1978, for an explicit attempt to make subjects aware of the representativeness and availability variables on which their stimuli varied.) In other studies, however, Tversky (1967, 1969, 1972a) has used schematic, decomposed stimuli in which the differences between objects (or profiles) were purposely made quite clear to the subject.

Social Judgment Theory

Early SJT studies often used real objects of considerable complexity (e.g., Hammond & Kern, 1959, had medical students judge various characteristics of 36 patients from sound motion pictures of patients being interviewed). In order to analyze the results of such studies, the complex stimuli were sometimes decomposed by having other subjects (sometimes experts) rate the real objects on a number of dimensions (see L. McClelland & Auslander, 1978, for a recent example of this technique).

In contrast, recent SJT studies (see Hammond, Rohrbaugh, Mumpower, & Adelman, 1977, for many examples) have used schematic stimuli almost exclusively. The subject in a typical experiment judges a series of "profiles" which are descriptions (usually numeric) of an object on a number of variables or cues. These profiles may be descriptions of either real or hypothetical objects. In selecting or constructing these profiles, SJT emphasizes that the set of schematic stimuli should be representative (in terms of formal properties such as means, ranges, and intercorrelations) of the real objects the subject is likely to encounter in the environment.

Information Integration Theory

The typical methods of IIT are difficult to place into one of our three categories. Many IIT studies have certainly used complex, real stimuli such as photographs (Lampel & Anderson, 1968; Shanteau & Nagy, 1976) and paragraph descriptions of events during the terms of each U. S. President (Anderson, 1973). However, the systematic

fashion (i.e., factorial design) in which such stimuli are combined with each other or with other stimulus components (e.g., trait words or probability phrases combined with the photographs) suggests that such stimuli are more like a priori schematic decompositions than whole, undecomposed objects. These IIT decompositions differ from the schematic profiles of other approaches in that only nominal levels are specified for each variable in contrast to the numerical profile descriptions used by DT and SJT. It should be noted that one of the advantages of IIT's functional measurement methodology is its ability to measure the subjective values of such stimuli without the necessity of prior numerical decomposition. Therefore, it seems that the methods of IIT are best described as being in between the whole object and schematic profile approaches.

Attribution Theory

Of all the six approaches, AT uses the most complex, undecomposed objects. The subjects often respond to elaborate scenarios (recorded on tape or acted live) in which the important variables are purposely not apparent. When schematic stimuli are used they are often like those used by IIT; that is, levels of relevant variables are defined nominally. For example, Weiner and Kukla (1970) defined student profiles in terms such as "has ability, did not expend effort, has a borderline test result," etc. and generally there are only two nominal levels for each variable (e.g., has ability vs. does not have ability).

Integration

The brief review above clearly demonstrates that all six approaches have sometimes used schematic, decomposed objects as stimuli. Such similarity across approaches should not be surprising for as many have noted, including Shanteau and Phelps (1977), "decomposition may well be the common foundation for judgment analysis." But "common" is still not "identical" and an important methodological difference does exist among these approaches. That is, IIT and AT use what we might label "minimal decompositions" while the other approaches favor (at least when they use schematic decomposition) "maximal decompositions." The IIT and AT decompositions are minimal in the sense that the stimulus presented to the subject still contains more or less real objects (photos, paragraphs) and specification of the levels of variables is at most nominal. Nominal levels are minimal because it is still the subjects' task to decide which level is more than another as well as how much more. For example in an IIT impression formation experiment the subject must decide, say, whether "methodical" or "organized" is the more likable trait.

In contrast, the maximal decompositions used in DT, and in some BDT, PDT, and SJT studies involve the presentation of relatively abstract stimuli which contain few if any components that could be considered as real objects. Instead, the schematic stimuli or profiles of these approaches describe objects mainly in terms of numerical values on specified variables, attributes, or cues. For example, in the Mexico City Airport study (Keeney,

1973a, and deNeufville & Keeney, 1972) describe alternative airport siting plans strictly in terms of numerical quantities such as cost, access time in minutes, number of aircraft operations per hour, number of people displaced by development, etc. There is clearly no ambiguity about whether one level of an attribute is "more than" another level (of course, the decision maker still decides which level is better).

We are not prepared to argue here whether minimal or maximal decomposition of the schematic stimuli is more appropriate for work in judgment and decision making. Most likely both are necessary in the complete approach and the choice depends on situational constraints as well as purposes for research or application. For example, minimal decomposition is obviously the preferred mode if the purpose of a basic research study is to discover empirically how people interpret information from real objects. On the other hand, if the purpose in an application were to prevent decision makers from using irrelevant or improper variables (e.g., "Is this potential airport site near to the property I own?") to evaluate alternatives, then maximal decomposition which makes explicit or eliminates such variables would be chosen. We suggest therefore that any integrated methodology for judgment and decision making must be able to employ both types of schematic decomposition.

While there seems to be agreement on the usefulness of some type of schematic decomposition, there is less agreement on the presentation of whole, nondecomposed objects or on the use of variables as stimuli. The choice to use decomposed vs. whole

stimulus-objects is closely related to the choice of idiographic or nomothetic methods. Approaches using real, complex stimuli (such as AT) tend to favor nomothetic methods because of the difficulty in constructing many stimuli while idiographic methods based on many judgments of different objects virtually require the use of easily constructable and repeatable stimuli.

[MORE]

Only BDT within the context of the SMART methodology (see Gardiner & Edwards, 1975, or Edwards, 1978) advocates the direct rating of variables in judgment and decision tasks. The claim is that the methods and procedures involved with presentation and analysis of whole objects or schematic stimuli are more complex than simply having the decision maker rank order and rate the importance of the key variables or attributes. Gardiner and Edwards argue that greater methodological complexity is not warranted because most of the models of judgment and decision making are insensitive to error in the estimation of parameters (but see McClelland, 1978, for a cautionary note about this supposed insensitivity).

E. Judgment Decomposition

Introduction

Whatever methodological choices are made with respect to stimulus-object decomposition, each approach must also decide how to decompose the judgments (or choices or preferences) during analysis. In general, all six approaches are remarkably similar with respect to judgment decomposition as all approaches analyze judgments in terms of weights (for each variable represented in the real or schematic stimulus) and function forms (which represent the relationship between objective levels of a variable and the person's judgments). (The names "weight" and "function form" are not used of course by all approaches but all approaches use concepts with essentially the same meaning; see the Theory section of this report.) However, the similarity ends there, for each approach uses quite distinct methods and procedures in order to decompose responses into weights and function forms. We briefly examine the methods of judgment decomposition employed by each approach.

Decision Theory

DT (in either its unidimensional or multiattribute forms) is based on a set of axioms which specify a set of conditions which must be satisfied before it makes sense to decompose a set of judgments into weights and function forms. Thus much of the DT data collection and analysis methods and procedures are devoted to testing those axioms. What to do when the axioms are violated is

not entirely clear; sometimes other variables or combination rules can be found which remove the violations while other times the decision maker is reeducated so that his judgments no longer violate the axioms. The important methodological point is that the decomposition does not proceed until it has been verified that the axioms are reasonably appropriate for the given context. Once the axioms are checked choice and indifference procedures (described later in the procedures section of this report) are used to determine the actual weights and function forms.

Behavioral Decision Theory

Gardiner and Edwards (1975) recognize that their SMART methods and procedures depend on several assumptions (axioms). For example, they discuss the "value independence" assumption but unlike DT do not consider its formal test in the data to be important. Indeed, in the data which they do collect (direct importance ratings of attributes) such tests would be impossible. They dismiss the need for such tests by appealing to the robustness of linear models and by criticizing the inconvenience in the methods and procedures of the other approaches. Because of this orientation BDT takes the most direct approach to assessing the weights and function forms--the judge directly estimates the weights on a ratio scale and linear function forms are imputed after the judge specifies minimum and maximum values and the direction (positive or negative) for each attribute.

Psychological Decision Theory

As yet prospect theory of the PDT approach has not been formally tested; it has only been proposed as an explanation for the systematic violations of SEU axioms which have been observed. Thus, although Kahneman and Tversky (1977) present an example of a decomposition (see their figure 4), it is not yet clear what methods should be used to decompose choices into the uncertainty weights and value functions which are the key components of prospect theory. Tversky (1967) used conjoint measurement to check the suitability as well as to produce the decompositions of his subjects' judgments of risky and riskless alternatives. Other works in the PDT tradition (e.g., Slovic & Lichtenstein) have used decomposition methods very similar to those of SJT which are described below.

Social Judgment Theory

Multiple regression is the decomposition method of SJT. Statistically determined beta weights estimate the judgment weights and function forms are derived by adding power terms (e.g., X^2) to the regression analysis. Unlike DT there are no axioms which must be satisfied in order to justify decomposition. Instead the adequacy of the decomposition is assessed post facto by examining the R^2 obtained from the regression. The advantage of this method over DT is that collection of data does not require that the judge satisfy any particular axioms. The disadvantage is that the interpretation of a low R^2 is ambiguous. A low value may be due either to inconsistency by the judge or to the inappropriateness of that

particular decomposition (e.g., use of wrong variables, combination rules, etc.). This ambiguity is recognized by SJT and methods for its resolution are indicated in Hammond, Stewart, Brehmer, and Steinmann (1975).

Information Integration Theory

The methods of IIT have in common with SJT the use of a statistical procedure for the decomposition of judgments into weights and scale values. (Note that the decomposition into scale values is slightly less restrictive than decomposition into function forms in that conceptual decomposition of the stimuli is not required prior to analysis; see the discussion of minimal and maximal schematic decomposition above.) The major difference is that IIT uses ANOVA instead of regression designs. That has the disadvantage of forcing the use of orthogonal sets of stimuli ("forcing" is too strong, but orthogonality is the usual practice because analysis of nonorthogonal designs can be extremely complex). The benefit is the ability to test statistically the "axioms" underlying various combination rules. For example, a significant "row by column" interaction invalidates an additive model and concentration of the interaction in the "linear by linear" or "bilinear" interaction term suggests a multiplicative model. The essence of the IIT methodology (functional measurement) is to check the validity of a specific algebraic model statistically and then to use that validated model to solve for the parameters necessary for the judgment decomposition.

Attribution Theory

AT is the only one of the six approaches which does not formally decompose judgments into weights and function forms. Because most AT experiments have only two levels for each variable, the concept of function form has no use. Although weights are not explicitly measured, the concept of relative weight is used implicitly in AT research. For example, results on situational versus dispositional attributions can be summarized by a statement such as "actors give more weight to situational variables in explaining their behavior while observers place more weight on dispositional factors for explaining the same behavior." But instead of decomposing judgments into weights and scale values, AT generally uses the method of diacritical confrontation (discussed in the theory section above) which pits two opposing hypotheses against each other in each judgment task. Thus, at most only relative weights are determined (e.g., "subjects give more weight to this than to that in making causal attributions"). It should be noted however that explicit measurement of weights would not be incompatible with most specific attribution theories--most particularly with Kelley's "ANOVA cube" model.

Integration

While the theoretical concepts for weights and function forms exist for prospect theory and attribution theory, the methods for accomplishing those decompositions have either not been developed or have not been adequately described to allow further discussion here.

We will therefore focus our comments in this section mostly on the methods of DT, BDT (specifically SMART), SJT, and IIT.

When applied to a policy problem all four approaches produce results (descriptions of judgment or preference policies in terms of weights and function forms) that are remarkably similar. The very different methods for arriving at such decompositions are summarized in Table III-E-1 in which the approaches are categorized according to their use of statistical analysis and formal tests of decomposition suitability.

By using statistical analysis (with error terms) IIT and SJT allow for a certain amount of inconsistency in the judgments whereas DT and SMART have no such mechanism for dealing with inconsistency. BDT simply claims that the robustness of the model obviates the problem. DT in its axiomatic form assumes no inconsistency (e.g., indifference judgments must be exact) so the DT methods and procedures use greater care in obtaining each judgment and choice (e.g., the decision analyst is almost always present to assist and check each judgment).

IIT and DT include formal checks for the suitability of decomposition and it is not performed unless the necessary conditions are found to be true. The methods of SJT and SMART do not include any similar assessments of the appropriateness of further analysis and decomposition. SJT does not perform such tests because it claims to be only a descriptive theory and the fact that all the judgments came from one individual is sufficient justification for describing those judgments in terms of weights and function forms.

TABLE III-E-1

CLASSIFICATION OF JUDGMENT DECOMPOSITION METHODS

FORMAL TESTS OF DECOMPOSITION SUITABILITY

| | | YES | NO |
|--|-----|-----|----------------|
| STATISTICAL ANALYSIS OF DECOMPOSITION | NO | DT | BDT (SMART) |
| | YES | IIT | SJT |

BDT admits the importance of the underlying axioms but omits their verification on the grounds that the axioms are generally true anyway (at least if the decision analyst is careful in developing the set of attributes) and because the typical model (the linear model) is very robust even when mild violations of the assumptions exist.

There is another basic difference among verification strategies which is not represented in Table III-E-1. DT and IIT seek justification for their decompositions in terms of internal properties of the data (e.g., independence, parallelism) while SJT looks more for external justification through an analysis of the judgment task. For example, if it were reasonable to conclude that all the judgments were from the same judge in the same state for the same task, then SJT would see no problem in proceeding with the decomposition; R^2 would then indicate the quality but not the appropriateness of the decomposition. DT and IIT would be less concerned with such task analysis as long as the necessary axioms and statistical conditions were true for the data set.

To further explicate the methodological differences and similarities among these approaches, we now examine how each of the four (DT, BDT-SMART, SJT, and IIT) would analyze a common data set. The hypothetical data set we shall use is based on a study of family composition preferences by Coombs, Coombs, & McClelland (1975). We choose this particular problem because as presented in its original form these judgments are not suitable for decomposition into weights and function forms. These data are listed in Table III-E-2 (the

Table III-E-2

Hypothetical Data for Comparing Decomposition Methods

| Attributes | | Judgment (Rating) | Rank Order (1 = Best) | Transformed Attributes | |
|------------|---|----------------------|--------------------------|---------------------------|----|
| X | Y | | | X* | Y* |
| 0 | 0 | 0 | 16 | 3 | 0 |
| 0 | 1 | 6 | 12 | 2 | 1 |
| 0 | 2 | 18 | 7 | 1 | 2 |
| 0 | 3 | 17 | 8 | 0 | 3 |
| 1 | 0 | 9 | 11 | 4 | 1 |
| 1 | 1 | 27 | 3 | 3 | 2 |
| 1 | 2 | 28 | 2 | 2 | 3 |
| 1 | 3 | 10 | 10 | 1 | 4 |
| 2 | 0 | 23 | 4 | 5 | 2 |
| 2 | 1 | 31 | 1 | 4 | 3 |
| 2 | 2 | 19 | 6 | 3 | 4 |
| 2 | 3 | 2 | 14 | 2 | 5 |
| 3 | 0 | 20 | 5 | 6 | 3 |
| 3 | 1 | 15 | 9 | 5 | 4 |
| 3 | 2 | 5 | 13 | 4 | 5 |
| 3 | 3 | 1 | 15 | 3 | 6 |

example is based on actual data with a few adjustments and assumptions added to make the data amenable to analysis by all four approaches). We now examine how each approach would detect and respond to the inappropriateness of judgment decomposition.

DT. Keeney and Raiffa would begin by checking the validity of the preferential independence axiom. The essence of that axiom is that if one prefers profile (x, y) to (x', y) then one should also prefer (x, y') to (x', y') no matter what value y' has. The point is that evaluation of x and x' should not depend on whether they are combined in the schematic stimulus with y or y' ; or in other words the function form for attribute X should not depend on the level of attribute Y . In this problem, Keeney and Raiffa would be likely (they have a number of alternate procedures so one cannot be stronger than "likely" but they do illustrate this procedure on pp. 200-300 of their 1976 book) to present the judge some of the following pairs and ask the judge to select the best or highest or whatever. Based on the "Rank Order" column of Table III-E-2 (with "1" being most preferred), our hypothetical decision maker's response is listed to the right of each pair.

| | |
|-------------------------------|------------------|
| Pair A: $(0, 0)$ vs. $(1, 0)$ | Choice: $(1, 0)$ |
| Pair B: $(0, 3)$ vs. $(1, 3)$ | Choice: $(0, 3)$ |
| Pair C: $(1, 1)$ vs. $(1, 2)$ | Choice: $(1, 2)$ |
| Pair D: $(2, 1)$ vs. $(2, 2)$ | Choice: $(2, 1)$ |

Normally more pairs would be presented than this, but these four would be sufficient to question the validity of preferential independence for these attributes and this decision maker.

From Pair A we would conclude that level 1 of attribute X is more valuable than level 0 since attribute Y is fixed. However, Pair B implies just the opposite conclusion ($X=0$ better than $X=1$) and thus the set of choices for those two pairs violates preferential independence. Pairs C and D demonstrate another violation of independence but this time with X fixed within each pair. Many other pairs of pairs from this data set also violate preferential independence. DT would therefore conclude that no decomposition into weights (they would call them scale factors) and function forms is feasible. The next step (we are now guessing somewhat) would be to explain the principle of preferential independence to the decision maker and to try to determine whether the violations were "real." The analysis would not proceed until either the decision maker changed his choices to conform to preferential independence or until the alternatives could be described on a set of attributes for which preferential independence was satisfied.

BDT (SMART). Gardiner and Edwards (1975) call the necessary decomposition property "value independence"--"the extent of your preference for property "value independence"--"the extent of your preference for location [x] over location [x'] of dimension [X] is unaffected by the position of the entity being evaluated on dimensions [Y, Z, ...]." However, in their method they have no way of checking

that property. In this case, they would presumably ask the decision maker to rate the relative importance of each attribute and to draw a value function for each. As we shall see later, while X and Y were reasonable attributes for describing the alternatives, they are not the appropriate attributes for decomposing the judgments. We are not sure, therefore, what would happen next. There are two possibilities: (a) the decision maker might throw up his hands and shout that the task made no sense to him because these are not the attributes he would use to make his decision or (b) he might be compliant and try to do the best he could in providing the necessary parameters. If he did the latter, then BDT might still detect the problem by comparing wholistic judgments with SMART-produced evaluations of the 16 alternatives.

Thus, we cannot be sure that SMART would detect the inappropriateness of further analysis nor is it clear what methods would be employed in case the impasse were discovered. Also note that monotonicity (either more of an attribute is better than less or vice-versa throughout its range) which is partly the basis for the robustness claim is clearly violated for these data.

PDT (Conjoint measurement). We have not explicitly considered conjoint measurement as a judgment and decision approach elsewhere in this report. However, it seems appropriate here to examine how conjoint methods would analyze our hypothetical data for two reasons: (a) Tversky whom we have identified as one of the key investigators within the PDT approach has also been one of the primary developers of the theory

and method of conjoint measurement (e.g., Krantz, Luce, Suppes, & Tversky, 1971; Krantz & Tversky, 1971) and (b) conjoint measurement was the method used in the study on which our example is based.

Conjoint measurement would check the independence or monotonicity axiom (identical to preferential independence in DT and value independence in BDT) and would find it seriously violated. The response of Coombs et al. to such violations was to determine whether the rank orderings and the axiom violations they implied were replicable; the thought was that maybe only the top two or three choices were really definite. However, the replication revealed that each subject produced essentially the same rank ordering as before so further analysis with conjoint measurement was not possible. This began a search for a transformation of the attributes for which independence was satisfied.

SJT. The decomposition method of SJT would be to enter these two attributes as independent variables and the judgment ratings as dependent variable in a multiple regression analysis. POLICY 2 (a special program which does the regression and then algebraically determines an appropriate weight-function form decomposition) yielded the results in Figure III-E-1. For these data and attributes, R^2 is only .43. SJT would regard this as an unacceptably low value. A likely followup would be to present replications to the judge (just as in the conjoint measurement approach) to determine the consistency of the judgments. If consistency were

RELATIVE WEIGHT PROFILE

| 0.0.....0.5.....1.0 | WEIGHT | FUNCT FORM |
|-------------------------------|--------|------------|
| CUE:1 AAAAAAAAAAAAAAAAAAAA | .43 | NONLIN |
| CUE:2 AAAAAAAAAAAAAAAAAAAA | .57 | NONLIN |
| 0.0.....0.5.....1.0 | | |

FUNCTION FORM PROFILE

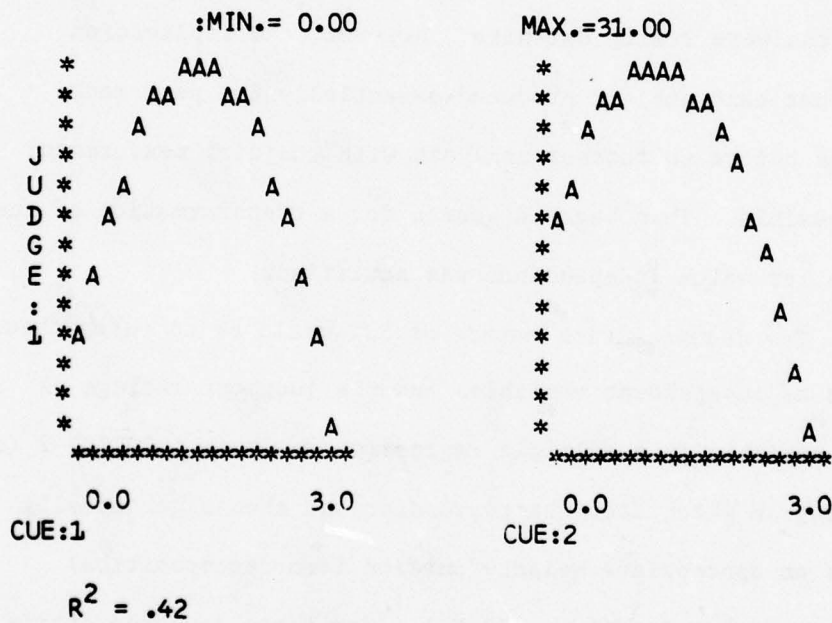


Figure III-E-1

Analysis of Hypothetical Data by SJT (Output from POLICY 2)

confirmed (as it would be in this case), then SJT would search for transformations of the attributes which would improve R^2 (although there seems to be little precedence for such a strategy within SJT).

IIT. Because there is only one subject and there are no replications in our data set, the standard IIT functional measurement methodology of using ANOVA would not be possible. Instead, IIT would most likely turn to a graphical test of the parallelism property required of any additive model. Parallelism is essentially equivalent to value and preferential independence (the shape of the function form for one attribute must not depend on the levels of the other attributes). Figure III-E-2 (a and b) shows that parallelism is clearly violated for these data. Each curve represents the function form for one attribute for each possible level of the other. A minimal requirement for parallelism is that the function forms should not intersect; the numerous crossings in these figures rule out constant function forms. IIT would view these figures as evidence of configurality and would not proceed further with the decomposition into weights and scale values. What other algebraic models would be tried next is not clear.

The transformation. Coombs, Coombs, and McClelland (1975) discovered that an appropriate transformation of attributes for these data is $X^* = X + Y$ and $Y^* = X - Y + 3$. Once this transformation is made (see Table III-E-2 again) all the problems blocking the decomposition methods

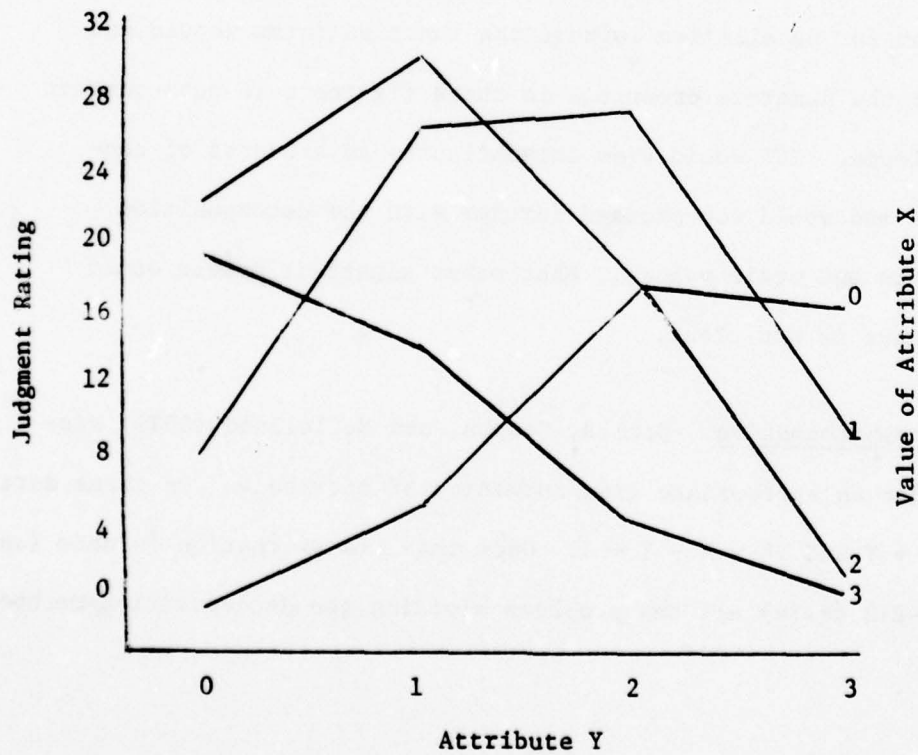
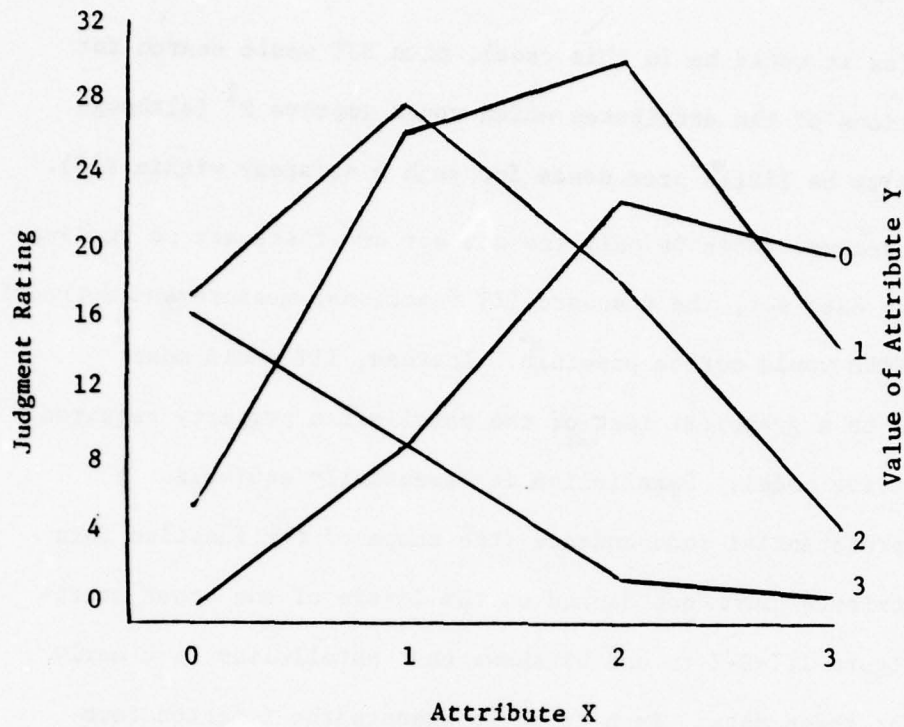


Figure III-E-2

IIT's Parallelism Test for the Hypothetical Data Set

of each approach are now resolved. For DT preferential independence is now satisfied [e.g., (5, 2) is preferred to (5, 4) and (1, 2) is preferred to (1, 4)]. So value independence is also satisfied for SMART and the decision maker is now able to draw a value function for each attribute (but it is still not monotone in the sense that more is always either better or worse). The function forms derived by SJT decomposition are displayed in Figure III-E-3; the respective relative weights are .26 and .74. R^2 is improved but still is only .78. (Note: a decomposition for these data does exist for which $R^2 = 1$; POLICY 2 failed to find it because the actual function forms are more complicated than the parabolas which the program tries to fit to the judgments.) The conjoint measurement analysis now proceeds smoothly (see Coombs et al. for details as well as the nomothetic evidence that this transformation is not a fluke). The parallelism test of IIT is graphed in Figure III-E-4 (because of the nonorthogonality introduced by the transformation only even-numbered attribute levels should be compared to other even-numbered levels and similarly for odd-numbered levels) and is clearly satisfied. Assuming equal weights, the scale values such as those listed in Table III-E-3 are derived algebraically by IIT. In summary, analysis now goes reasonably well for all approaches.

So what have we learned from this rather extended example? We think that there are at least two very important conclusions which we consider in turn below.

First, this example demonstrates that despite the very different methods used to check model validity (e.g., axioms vs. statistics)

RELATIVE WEIGHT PROFILE

| 0.0.....0.5.....1.0 | WEIGHT | FUNCT FORM |
|---------------------------------------|--------|------------|
| CUE:1 AAAAAAAAAAAA | .26 | NONLIN |
| CUE:2 AAAAAAAAAAAAAAAAAAAAAAAAAAAA | .74 | NONLIN |
| 0.0.....0.5.....1.0 | | |

FUNCTION FORM PROFILE

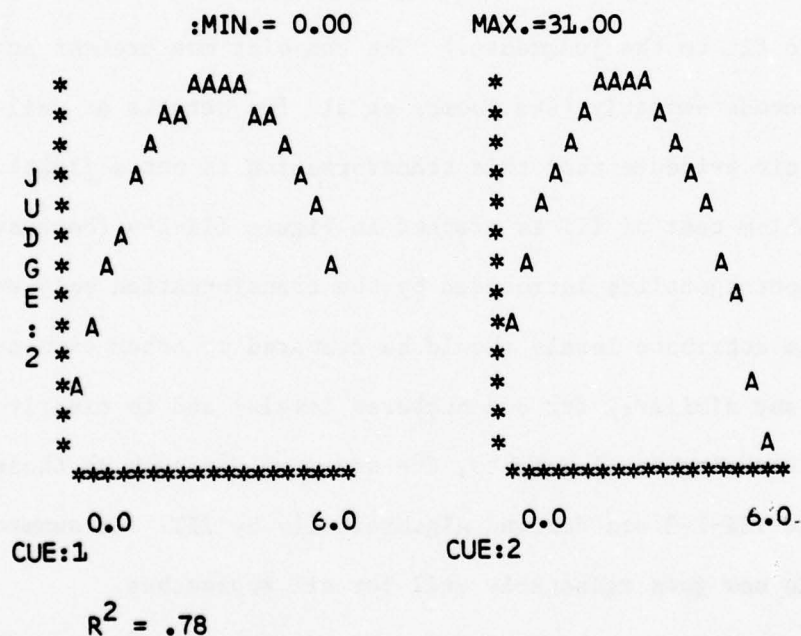


Figure III-E-3

SJT Analysis of Transformed Data

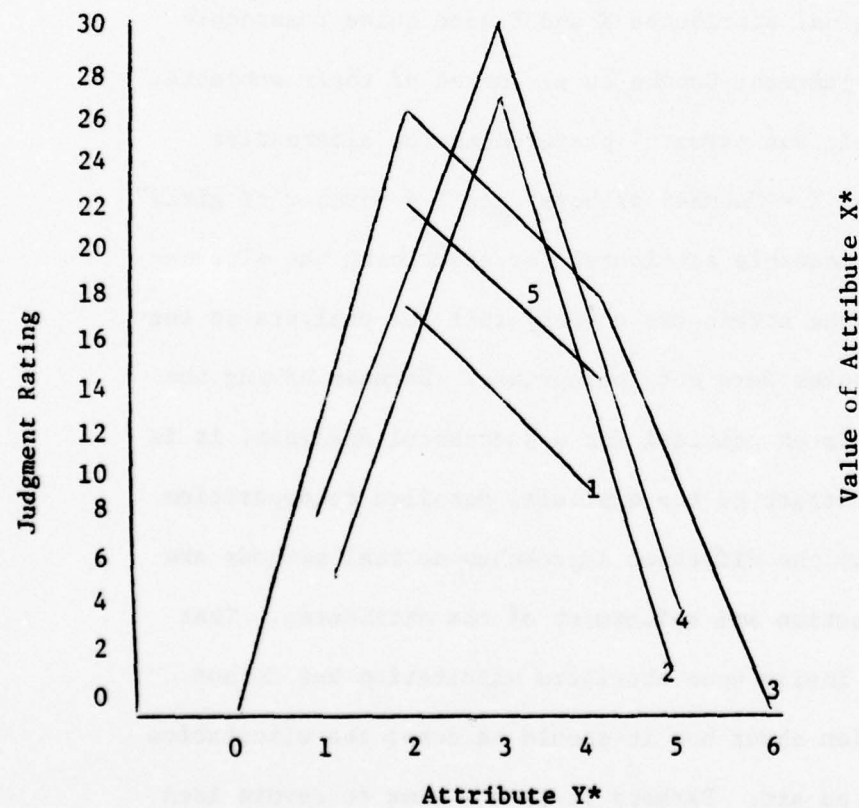
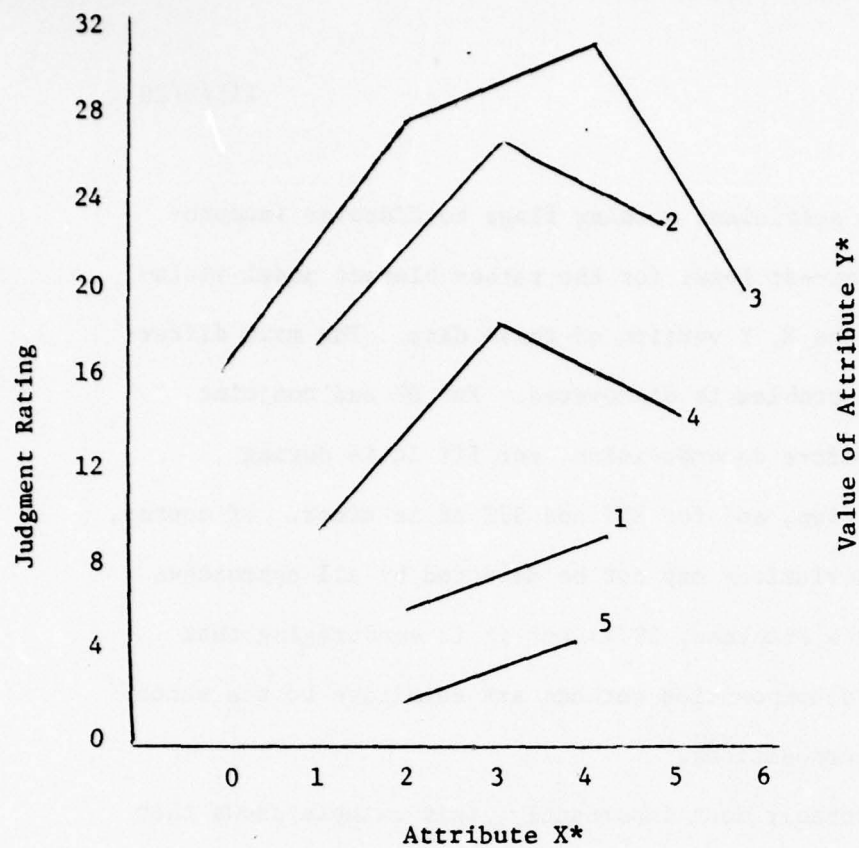


Figure III-E-4

IIT's Parallelism Test for Transformed Attributes

all approaches have sufficient warning flags to discover inappropriate decompositions--at least for the rather blatant model violations contained in the X, Y version of these data. The main difference is in when the problem is discovered. For DT and conjoint measurement it is before decomposition, for IIT it is during attempted decomposition, and for BDT and SJT it is after. Of course, more subtle model deviations may not be detected by all approaches (e.g., see Anderson & Shanteau, 1977) but it is encouraging that the very different decomposition methods are sensitive to the suitability of those decompositions.

Second, and probably most importantly, this example shows that having the correct attributes is essential for a successful decomposition. The original attributes X and Y were quite reasonable for the preference judgment Coombs et al. asked of their subjects. The substantive topic was parents' preferences for alternative family compositions; X = "number of boys" and Y = "number of girls" seemed perfectly reasonable attributes for describing the alternatives. However, those attributes clearly fail for analysis so the "reasonable" attributes were not appropriate. Because having the correct attributes is so critical for a successful analysis, it is striking that in contrast to the explicit, detailed decomposition methods specified by the different approaches no real methods are provided for elicitation and refinement of the attributes. That is, all approaches insist upon attribute elicitation but do not give much instruction about how it should be done; the elicitation process is left as an art. Perhaps it is now time to devote less

attention to developing more and more refined decomposition methods within approaches in order to devote greater attention to the elicitation problem--since it afflicts all approaches.

[MORE]

For this particular example in which the correct attributes were related by a simple transformation (i.e., a 45° axis rotation) to the original attributes, it is interesting to ask whether any of the various decomposition methods would have had any advantage over the others in finding such a transformation. Of course we know that investigators using conjoint measurement did find the appropriate transformation. However, there does not seem to have been anything specific in the methodology of conjoint measurement which aided Coombs et al. in their search. They simply generated a number of plausible decision rules which their subjects might have used and then tested the implications of those rules in the data. Any method which allows the testing of alternative composition rules or organizing principles (i.e., DT, SJT, or IIT but not SMART or BDT) could have done much the same. Raiffa (1968, p. 258) did consider precisely this transformation as a means of achieving judgmental independence for certain problems so DT could claim that it is at least aware of the need to look for such transformations. But does DT have any explicit methods for the search? Unfortunately not. Raiffa (1968, p. 259) states:

We can generalize and say that it is sometimes possible to circumvent the problem of judgmental dependence by concocting meaningful, auxiliary, mathematically related quantities that are judgmentally independent. This is one of the tricks of the trade. (*italics ours*)

Indeed, all approaches seem to have left the problems of eliciting, defining, and transforming attributes as "tricks of the trade." Again, this stands in sharp contrast to the extremely explicit methods advocated for using those attributes once obtained.

Our first attempt to apply the various methods to a common data set has convinced us that such a procedure is an invaluable tool for highlighting methodological similarities, differences, and gaps among the approaches. Other attempts to apply the methods of the approaches to a common problem (e.g., Shanteau & Phelps, 1977) have also been extremely informative so we can only conclude that further integration of judgment and decision approaches will require many more such efforts. There are certainly many more issues to be explored, for example, the difference in IIT's and SJT's methodological uses of the weight parameters. To be really successful however, it is probably necessary that these applications of different methods to the same problem be done by representatives from each approach.

[MORE]

F. Methods of Partitioning the Decision

Introduction

As a result of attempts to make complex decision and judgment problems tractable both for decision makers and for research, DT, BDT, and SJT have developed explicit methods by which the decision maker can "divide and conquer" the decision or policy problem. The general methods used by the various approaches are so similar that separate descriptions of each would be redundant. Therefore, communalities and differences between the two basic methods are noted first.

Methods for complex decisions. Figure III-F-1 depicts the prototypical decision or policy problem. The adjective "prototypical" is justifiable because more or less the same figure appears in publications from DT (see Figure 1 in Keeney, 1976, and Figure 7.6 in Keeney & Raiffa, 1976), BDT (Figure 3 in Gardiner & Edwards, 1975) and SJT (Figure 2 in Hammond, Rohrbaugh, Mumpower, & Adelman, 1977). As indicated in Figure III-F-1, there are two basic problems in solving any decision or policy problem. The first is the "knowledge" problem and involves identifying the options and then describing those options on variables (dimensions, cues, attributes, etc.) upon which a meaningful policy can be based. For example, when Hammond and Adelman (1976) helped a city council select handgun ammunition for a police department the bullets were described in terms of injury,

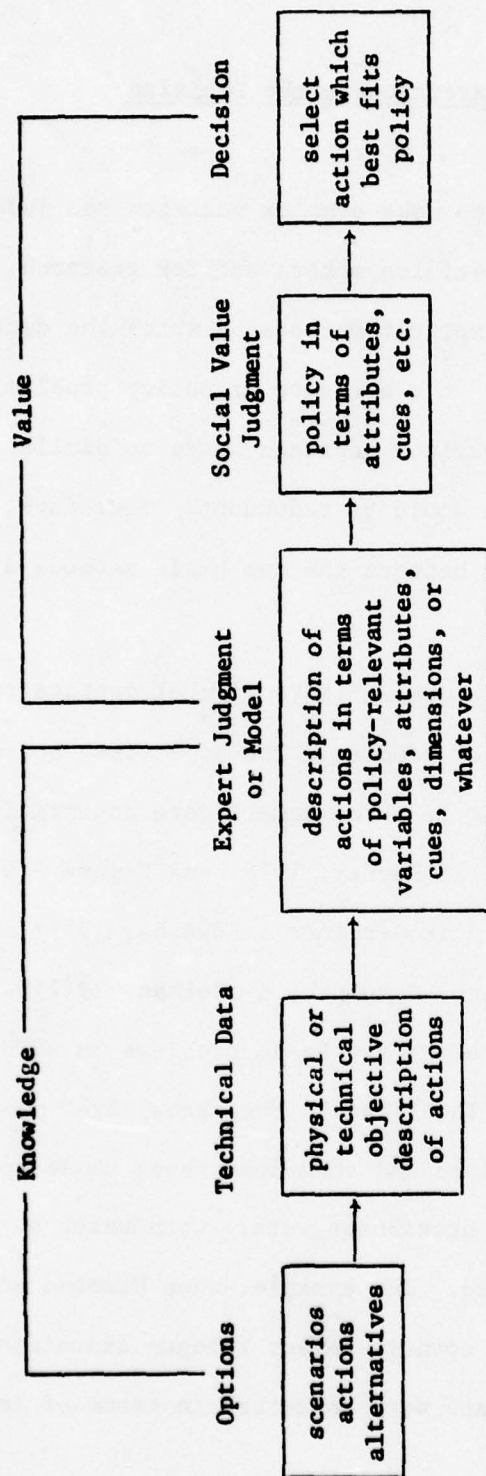


Figure III-F-1. Diagram of Prototypical Social Value Problem

stopping effectiveness, and danger to bystanders--attributes that were meaningful for policy judgments but far removed from the physical properties of the original bullets. Second is the "value" problem, involving a description of the system that is to be used to evaluate or judge an option described on the variables resulting from the solution to the "knowledge" part of the problem. The value section culminates in the selection of the option that best "fits" (usually maximizes but there can be other definitions of fit) the policy or the decision maker's values.

The differences between the knowledge and value problems in Figure III-F-1 correspond to the different interests of the Group I and Group II approaches as they were broadly defined in the Theory section. Historically, Group II psychological theories of judgment have aimed at the knowledge problem (they have been more than that of course) and judgment for choice or decision has been treated as a special case if at all. Not only is this a true statement for SJT, IIT, and AT but also for other psychological approaches to judgment. The key problem in the knowledge section is locating an entity on various judgment dimensions and certainly all the research on clinical judgment fits into that paradigm. But until recently most judgment theories have had little interest in the value portion of the figure. As a consequence most of the methods of psychological approaches were developed to describe, explain, or predict knowledge that the judge had (or thought he had).

On the other hand, Group I approaches have concentrated traditionally on the value questions, giving only scant attention to the

knowledge questions. So while elaborate methods have been developed to measure values (utilities) not that much attention has been given to measurement of "facts." "Knowledge" has entered DT usually only in the form of subjective probabilities. Because these probabilities were subjective there was usually no attempt to uncover the bases of those probabilities; they were just assessed and taken as is as long as they were consistent. One exception is the use of Bayesian statistics to incorporate new information, but that is really only an organizing principle and still says nothing about the origin of the subjective probabilities.

The above stereotypes are too pat but probably not all that unfair. However, while those stereotypes of basic judgment and decision methods apply to the historical or traditional approaches, the distinction has become increasingly blurred since about 1970 (we have no idea why 1970 seems to be such a watershed year). Raiffa's first MAUT paper was in 1969 and the first actual applications occurred in the early 1970s; Edwards' social value paper appeared in 1971; SJT social policy applications also began in the early 1970s and the 1975 "bullet" study seems to have been SJT's first full-scale application involving both the knowledge and value problems; although IIT has no practical applications to social policy the joint work of Shanteau and Anderson around 1970 moved IIT into the value portion of Figure III-F-1. Because both approaches were now coping with the total decision or policy problem, judgment approaches had to deal with values and decision approaches and had to consider the knowledge issues more seriously.

So what did each basic approach do when faced with new, unfamiliar problems to solve? Each simply transferred as much of their existing methods for one domain (knowledge or value) into the other. We follow below in more detail exactly what each approach did as it entered new territory.

Decision Theory

The essence of the knowledge problem is describing the alternatives on attributes upon which it makes sense to have preferences. Ideally, DT would like to have a probability distribution for each option on each attribute but DT often settles for point-estimates in most applications. And how does DT get those probability distributions or point-estimates? Any way possible. Keeney and Sicherman (1976) clearly state the eclectic approach of DT to solving the knowledge problem: The probabilities "may be specified using any combination of analytical models, simulation models, subjective assessments, and data that is available and appropriate." Common methods are the use of simulation models and expert judgment; both are illustrated in a study of fire department operations (Keeney, 1973b; see also Keeney & Raiffa, 1976, Chapter 7) which we describe below.

The problem is to form a policy concerning quantity and quality of fire equipment, location of equipment, quantity and quality of personnel, etc.--the goal being the most efficient use of available resources. But these are really just input variables that are not very meaningful for describing a policy or measuring fire department

efficiency. So first (see Figure III-F-2 which is reproduced from Keeney & Raiffa, 1976) a simulation model developed from past objective data translates locations on those input variables into the somewhat more meaningful attributes of engine and ladder response times (firemen supposedly evaluate everything in terms of response time). That still does not solve the knowledge problem, for what the decision maker needs to know is how each policy option affects number of fatalities and injuries, amount of physical damage, physical and mental fatigue of fire department personnel, etc.

For various reasons it was not possible to use the objective data to relate response time to the policy attributes and no simulation model existed for this purpose so Keeney turned to expert judgment to supply the missing link. The important methodological point is that choice procedures were used to assess the technical expert's knowledge. In this case, Keeney measured the deputy fire chief's preferences for various engine and ladder response times. Standard DT assessment techniques involving lotteries were used. For example, the chief was indifferent between a 50-50 chance that the first ladder would arrive in either 1 or 5 minutes or that the first ladder would arrive in 3.4 minutes for sure.

Interestingly the alternatives are never described in terms of the final evaluation attributes (fatalities, etc.). Instead the deputy chief's preferences are substituted as a "proxy" attribute with the presumption that his preferences are based solely on his knowledge of how best to fight fires and his knowledge of what equipment is needed when. This proxy attribute method has the

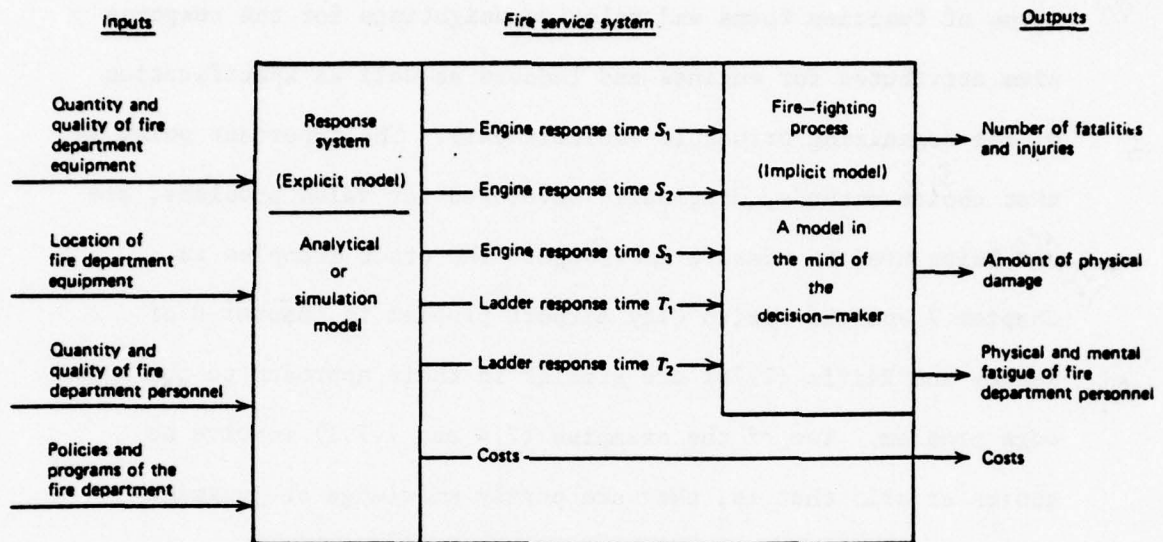


Figure III-F-2. A simplified model of a fire department service system.
(From Keeney & Raiffa, 1976, p. 379)

effect of hiding his implicit weighting of fatalities, property damage, fireman fatigue, etc. in his definition of the best way to fight fires. However not everything is hidden for the analysis produces a description of the chief's expert judgment policy in terms of function forms and relative weightings for the response time attributes for engines and ladders as well as specification of the organizing principle (multilinear). The important point is that choice methods, originally developed for value problems, are now being used to measure knowledge. The other examples in Chapter 7 and the Mexico City airport problem in Chapter 8 of Keeney and Raiffa (1976) are similar in their approach to the knowledge problem. Two of the examples (7.4 and 7.7.1) involve no choice at all; that is, they are purely knowledge or judgment studies!

Behavioral Decision Theory (SMART)

Step 8 of the SMART procedure (see Gardiner & Edwards, 1975) is "measuring of the location of each entity on each dimension" which we consider to be part of the knowledge problem. Unfortunately they do not have very explicit methods as to how this is to be done. They do distinguish three types of dimensions--(a) purely subjective, (b) partly subjective, and (c) purely objective. In the purely subjective case the knowledge problem is solved by "simply getting an appropriate expert to estimate the position of that entity on that dimension on a 0 to 100 scale." They do suggest that the knowledge problem is not all that difficult so that the complex procedures of DT are not required. For example,

they indicate that any disagreements at Step 8 (disagreements about fact) should be resolved by "the simple expedient of asking only the best available expert for each dimension to make (fact) judgments about that dimension." More recent work within BDT has gone beyond this to use more sophisticated solution methods such as simulation models, but in general the BDT methods pertain mostly to the value problem and leave the knowledge problem to someone else.

Social Judgment Theory

SJT is a cognitive theory and its methods pertain to the knowledge part of Figure III-F-2. Those methods have been aimed at assessing a person's knowledge in terms of utilization of cue information. In the double-system case it is even possible to assess the quality of a person's knowledge and it is always possible to measure the consistency of that knowledge (just as it was possible for Keeney to assess the chief's knowledge about the escalation of fires).

But what is to be done when policy choices must be made? The transition is easy: value judgments are just another type of judgment and the same methodology can be applied. However, for value judgments we are always limited to the single-system case and many nice features of the lens model disappear. One disadvantage is that a legacy of the knowledge approach is often transferred to the value arena. For example, when Hammond, Mumpower, and Smith (1977) ask the reader to "consider a person who must form

a judgment about the type of growth appropriate for his community [emphasis added]" it almost seems as if there is a "right" answer to be found. It is slightly more clear that values are involved when "The policy maker's task is. . .to integrate information. . . into a judgment that indicates a preference [emphasis added]."

Notice the reversal from DT: for SJT judgments indicate preferences but for DT preferences indicate judgments! So just as DT moved its choice methods from the value part to the knowledge part of the problem, SJT has moved its judgment methods from the knowledge part to the value part. (This is clearly illustrated in Hammond & Adelman, 1976.)

Integration

For the three approaches which have been applied to social problems (DT, BDT, and SJT), a degree of integration already exists--all would agree that Figure III-F-1 more or less represents the basic structure of policy decisions. An important strategy common to all three approaches is the methodological separation of the knowledge and value problems of Figure III-F-1. The separation is often so complete that the judgments (or choices) of different people are used to solve the knowledge and value problems. Such separation is a reflection of the "divide and conquer" strategy which permeates most approaches to judgment and decision making. Hammond and Adelman (1976) also claim an additional advantage: "The separation phase permits elected representatives to function exclusively as policy-makers, and scientists to function exclusively as scientists."

While there is agreement on the global methods for policy and decision problems, we have noted that each approach has had to apply its methods to a new domain in order to solve the complete problem. None of the approaches seems to have given much thought as to whether its transfers are appropriate or whether other better alternatives exist. One obvious possibility would be to use judgment methods for knowledge problems and choice methods for value problems. Just as the proponents of the various approaches, we have yet to do the necessary thinking to answer such questions.

[MORE]

In the most recent applications of their methods, DT, BDT, and SJT all seem to be turning whenever possible to the use of simulation models to replace the judgments of technical experts for solving the knowledge problem. With more and more reliance on computer models the interesting question arises as to how, if at all, human judgment influences the development of such models. Are the judgment aspects hidden in those models? How do such models get built and do any of the six approaches have any suggestions for improving them or facilitating their construction?

[MORE]

There are, of course, many other methods that have been proposed for aiding decision makers in complex judgment and decision tasks. Two important classes of such methods are debiasing techniques and consistency-improving techniques. "Debiasing" methods are based on the premise that many "errors of judgment are often systematic rather than random, manifesting bias rather than

confusion" (Kahneman & Tversky, 1978). Conversely, consistency-improving techniques are aimed at reducing those random errors of judgment that are due to confusion and inconsistency. We had intended to describe and discuss details of both types of methods, but we have simply run out of time. We refer the reader to Kahneman and Tversky (1978), Lichtenstein, Slovic, Fischhoff, Layman, and Combs (1978), and Timmers and Wagenaar (1977) for examples of debiasing methods and to Goldberg (1970, 1976), Dawes (1971), Dawes and Corrigan (1974), and Hammond, Rohrbaugh, Mumpower, and Adelman (1977) for consistency techniques.

Our perusal of the debiasing and consistency methods suggested to us that an important integrative step in this area would be to catalog the many judgment and decision problems and their proposed remedies. Such a catalog might form the basis of a "Decision and Judgment Analyst's Reference" somewhat analogous to the well-known Physician's Desk Reference. This reference would list specific judgment diseases ("diseases" is too strong because many of the biases and errors of judgment seem to result from normal, adaptive heuristics that cause trouble only in certain situations but "disease" will do for now), the signs and symptoms useful for diagnosing the judgment diseases, the presumed causes, and finally proposed remedies (with "clinical trial" information for the remedy). Tables III-F-1 to III-F-3 illustrate sample entries in such a catalog.

Table III-F-1

Sample Entry in "Decision Analyst's Reference"

Disease: Misperception (underestimation) of exponential growth.

Symptoms or Signs:

- 1) severe underestimation of level that a growth process will reach at a specific future time.
- 2) severe underestimation of time before a growth process will reach a critical level.

Causes:

- 1) use of successive differences (as opposed to successive ratios) to estimate future numbers given initial sequence
- 2) anchoring

Remedies: change estimation problem so that the growth process is decreasing (i.e., inverse exponential growth).

Example: original problem is to estimate population of county X; change problem so that task is to estimate square miles per inhabitant. Population is increasing growth process while square miles per inhabitant is decreasing.

Validation of Remedy: Timmers and Wagenaar (1977)

Plausible remedies known to be ineffective:

- 1) graphical presentation of growth process only increases degree of underestimation.
- 2) more detailed and frequent information about growth process procedures greater underestimation (Note: this implies that "experts" very familiar with past history of growth process

Table III-F-1 continued.

may give the worst estimates. For example, a former resident who returns to county X after an absence of 10 years may better estimate growth rate than a resident who has lived continually in the county)

- 3) obtaining parameter (starting level, growth exponent, etc.) estimates to be aggregated mechanically. (Wagenaar demonstrates that the disease is due mainly to misperception rather than misaggregation. Note that this may be unique among judgment and decision diseases.)

Table III-F-2

Sample Entry in "Decision Analyst's Reference"

Disease: Non-regressing predictions for unknown quantities.

Symptoms or Signs: predicted value is selected so that the standing of the case in the distribution of outcomes matches its standing in the distribution of impressions.

Causes:

- 1) failure to take uncertainty into account
- 2) insensitivity to reliability and predictive validity of available information
- 3) assumption that deviations from predictions are equally likely to be toward the average as away from it.

Remedies: (Kahneman & Tversky, 1978)

Step 1: Select reference class that can provide distributional information.

Step 2: Assess distribution for the reference class either from past data or from expert estimates of mean (μ) and variability (σ^2).

Step 3: Expert makes non-regressive intuitive estimate y_{est} on the basis of singular (nondistributed) information about the particular case.

Step 4: assess predictability of available information either from past performance data or by expert's estimate (ρ).

Table III-F-2 continued.

Step 5: Analytically combine the information from Steps 2,
3, and 4 according to the following formula:

$$\hat{y} = \mu + \rho(y_{\text{est}} - \mu)$$

Step 6: Explain rational of computation to expert, redo Steps
2 - 5 if expert wishes.

Experimental Validation of Remedy: none

Table III-F-3

Sample Entry in "Decision Analyst's Reference"

Disease: Overconfidence in the estimates of unknown quantities.

Symptoms or Signs: too many "surprises" when using subjective confidence intervals.

(Note: overconfidence is not expected to occur in essentially repetitive situations--weather forecasting, etc.).

Causes:

- 1) insensitivity to factors that determine the quality of evidence (amount, reliability, and sample size of available information).
- 2) oversensitivity to the consistency of available data. This is further exacerbated by a tendency to overestimate consistency ("illusory correlation").
- 3) making estimates conditional on (often unstated) restrictive assumptions (e.g., "under normal operating conditions").
- 4) anchoring

Remedies: (Kahneman & Tversky, 1978)

Use Steps 1 - 5 for correcting non-regressive predictions.

Step 6: Analytically combine the information from Steps 2, 3, 4, and 5 according to the following formula:

Confidence interval: $\mu \pm \frac{c}{2} \sqrt{1 - p^2}$ where c is the width for the corresponding confidence interval for the reference class distribution.

Experimental Validation of Remedy: none.

PROCEDURE

Introduction

This section discusses the procedures of the six approaches. The appropriate distinction between methods (discussed in the previous section) and procedures is sometimes unclear. Most dictionaries, for example, provide almost synonymous definitions of the two terms. We have attempted in this report to maintain a distinction that is similar to that between strategy and tactics. The methodology section was concerned with "grand plans" for action, while this section is concerned with particular techniques used to carry out those plans.

Our attempts to contrast and compare procedures across approaches encountered several difficulties deserving mention. First, while proponents' descriptions of their approaches usually discuss at length their theoretical starting points and principles and frequently describe in detail their methodological orientation, neither specific procedures for implementation nor the links between theory, method, and procedure are commonly found in such descriptions. For most of the approaches, there is no one source that provides a definitive and exhaustive description of its procedures. Since the authors had little or no direct experience with the procedures of several of the approaches, we therefore had to rely substantially upon review and interpretation of multiple sources in developing our descriptions of the approaches' procedures. Obviously such an approach is less than ideal and likely to lead to incompleteness, inaccuracy, and imprecision. We apologize for such shortcomings and request the reader's help in redressing them.

A second problem was that there is ordinarily more than one set of procedures by which an approach's aims can be carried out. For almost any generalization concerning an approach's procedures, an exception can somewhere be found. If consistency is the hobgoblin of small minds, then, (if one looks only to the procedures they use) there are only superior intellects at work on judgment and decision problems. Moreover, the situation is complicated by the fact that various approaches have different intended aims and functions. Some approaches are concerned almost exclusively with basic research on judgment and decision processes, some are oriented more toward practical applications, and still others share both aims (see the Intended Function section above). And, for those approaches with multiple aims and functions, the procedures used in basic research are not always identical to those used in practical application. (The difficulties in this respect were made salient when the authors found that, not infrequently, we disagreed among ourselves concerning the procedures of those approaches with which we considered ourselves quite familiar.)

Finally, the appropriate criteria for comparing and contrasting procedures are not readily apparent. For example, the major criterion might conceivably have been congruence with theoretical precepts and methodological principles. Such congruence would seem to possess considerable validity as a criterion, since coherence among theory, method, and procedure seems necessary for an approach to be described as well-developed. On the other hand, an argument could be made that the most appropriate criteria for comparison and

contrast would be pragmatic in nature. That is, it could be argued that the comparison should focus on such questions as, say, which procedures are easiest to use or which lead to the most "veridical" descriptions of judgment and decision processes. And, additional candidates for appropriate criteria are easily generated.

Our decision in response to these considerations was to make this section of the report as descriptive, neutral, and nonevaluative as possible. We will attempt simply to identify and describe as thoroughly as possible the similarities and differences among approaches in their procedures. We will try to describe the most common procedures in use within an approach. We will consider both the procedures used in basic research and the procedures used in practical application for those approaches sharing both intended functions. We will try not to speculate about the implications of similarities and differences among approaches in procedures.

The discussion will be organized around a slightly revised version of the lens model diagram presented in Figure II-F-1 of the theory section; i.e., procedures are discussed according to their location with respect to Figure IV-1. The framework of the discussion is further defined by a series of summary tables (Tables IV-A-G) corresponding to each of the seven major loci of Figure IV-1. Those loci are (a) task construction and cue selection, (b) task characteristics (both objective and subjective), (c) the subjective use of subjective data, (d) the subjective use of objective data, (e) cue utilization, (f) central processes, and (g) task-judge interaction. The tables identify major procedural points and major concepts associated with each loci.

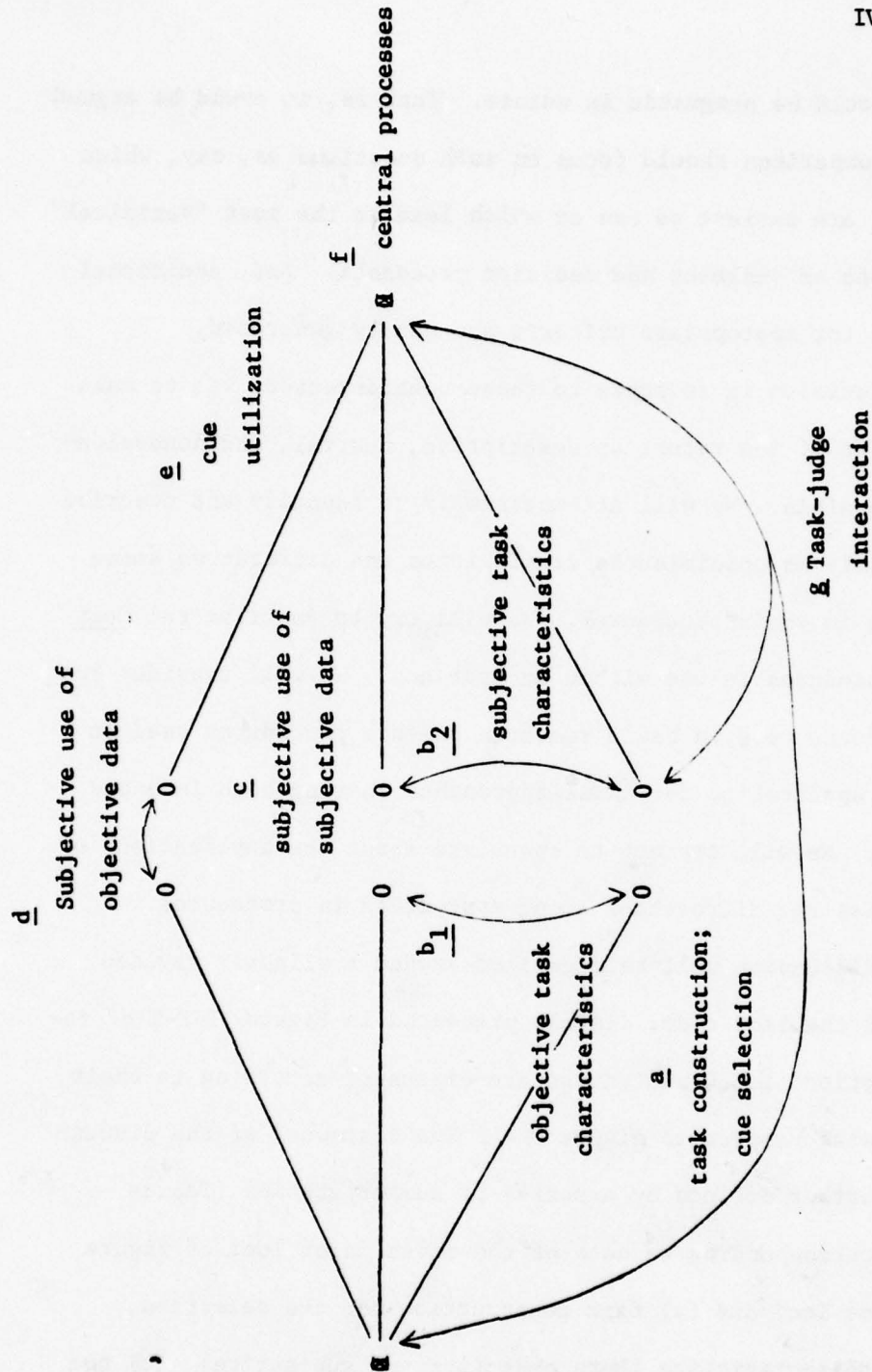


Figure IV-1
(Version of Figure II-F-1)

The narratives accompanying the tables describe the manner in which approaches differ and are similar with respect to the procedures at each of the seven loci. In addition, they describe the manner in which each approach defines or operationalizes similar or identical major concepts. The narratives also try to relate differences in procedures among approaches to differences in theory or method. Each narrative concludes with a brief summary or overview.

A. Task Construction/Cue Selection (Table IV-A)

A fitting place to start a discussion of procedures is with task construction and cue selection. The issue here is the manner in which the judgment or decision analyst proceeds in constructing and presenting a task for the judge or decision maker.

Specification by analyst versus judge. A primary question is that of who defines the nature of the judgment or decision problem and its dimensions. For three of the approaches the answer is sometimes or usually the judge or decision maker himself.

DT's exclusive intended function is to aid decision makers with applied decision problems. Without exception, therefore, the decision maker defines the problems and its dimensions, although with the help and direction of the decision analyst. BDT and SJT, in their applied versions, are also intended to aid decision makers in applied decision problems. In such applications, the decision maker defines the problem and its dimensions, again, with considerable help and direction from the analyst.

All three approaches are somewhat imprecise concerning their procedures for the important step of eliciting specification of the

Table IV-A
Task Construction/Cue Selection^a

| DT | BDT | PDT | SJT | IIT | AT |
|--|-----|-----|-------------------|-----|----|
| <u>Task construction - Specification by analyst vs. judge</u> | | | | | |
| Definition of the nature of the problem and its dimensions by decision maker | X | | X | | |
| Definition of problem by researcher | | X | X | X | X |
| <u>Task construction - Fact/knowledge distinction</u> | | | | | |
| Distinction maintained | X | | X | | |
| <u>Task construction - Type of response (Preference/Utility)</u> | | | | | |
| Lotteries | X | X | | | |
| Direct Specification | | | X | | |
| | | X | Wholistic Ratings | X | |
| <u>Task construction - Type of response (Inference)</u> | | | | | |
| Choice | | X | | | X |
| | | X | Evaluations | X | |

Table IV-A continued

| DT | BDT | PDT | SJT | IIT | AT |
|--|--------------|-----|-------------------------|--------------|----|
| <u>Task construction - Type of response (Probability estimation)</u> | | | | | |
| Lotteries | X | | | | |
| Direct Estimation | | X | | | |
| <u>Task construction - Design</u> | | | | | |
| X | (no objects) | X | (Regression) | WS-Factorial | X |
| | BS-Factorial | X | X | X | |
| <u>Cue selection - Level of abstraction</u> | | | | | |
| Decomposed objects | X | | X | X | X |
| Variables | | | | | |
| <u>Cue selection - Metric</u> | | | | | |
| X | X | | Quantitative-continuous | X | |
| Qualitative Discrete | X | X | X | X | X |

^aX's indicate that the approaches so indicated share this point or concept with the approach under which the point or concept appears.

problem and its dimensions from the decision maker. All three view the process as a creative one, which is much more an art than a science. There is little likelihood, for example, that proponents of BDT or SJT would disagree with Keeney and Raiffa (1976) in their exposition on DT, when they say:

"The intertwined processes of articulating objectives and identifying attributes are basically creative in nature. Thus, it is not possible to establish a step-by-step procedure that leads one in the end to a meaningful set of objectives and attributes" (p. 64).

Of the three approaches, DT discusses most extensively procedures for specifying the decision problem and its dimensions (e.g., Keeney & Raiffa, 1976, pp. 31-65). They recommend first that objectives be generated for specific problems. Objectives may be identified by examination of the relevant literature and analytical study of the problem, as well as by interaction with the decision maker(s). Questions put to decision makers are of the sort, "What's important?", "Why is it important?", "What else is important?", etc.

The next step is the location of identified objectives into a hierarchy. The hierarchy establishes the relative importance and generality of objectives. Major objectives are more general and more inclusive than the lower order objectives related to them. The hierarchy is usually established by a trial-and-error procedure involving questions about mean-ends relationships. The most general test for whether or not an objective should be included in the hierarchy is whether the decision maker(s) answers affirmatively

when asked whether the best imaginable course of action might change if that objective were excluded from the hierarchy.

Finally, attributes associated with each objective must be identified. Attributes should be comprehensive and measurable. An attribute is comprehensive if the particular levels of an attribute clearly indicate the extent to which the associated objective has been achieved. An attribute is measurable if it is feasible to obtain estimates of (a) the probability distribution for each alternative over the possible levels of the attribute and (b) the decision maker's preferences for different levels of the attributes. Some attributes are objective, measuring the associated objective directly (e.g., the attribute dollars may be associated with the objective "increase profits"). They may also be subjective indices (e.g., measured (subjectively), proxy attributes (e.g., which measure the objective indirectly), or direct preference measurements. Procedurally, attributes for association with specific objectives are selected on a trial-and-error basis, with the analyst asking questions to insure that the attributes are comprehensive and measurable (e.g., "Is there anything else you would need to know other than the value of attribute x in order to know how well objective X was achieved?"). The decision analyst also questions the decision maker and inspects the final set of attributes in order to test that the set is complete, operational, decomposable, nonredundant, and minimal.

BDT discusses the step of defining the decision problem and its dimensions less extensively than DT, but more so than SJT.

Although the procedures are described in less detail than by DT, they appear to be quite similar. Procedurally, BDT (Gardiner & Edwards, 1975) notes (a) that often it is practical and useful to ignore the hierarchical structure of goals or objectives and instead to specify a simple list of goals that seem important for the purpose at hand and (b) that the number of objectives or goals identified be kept to a minimum.

SJT does not discuss at length the appropriate procedures for identifying and defining the judgment or decision problems, noting (Hammond, Stewart, Brehmer, & Steinmann, 1975) that the "methods used in this step are highly situation (and investigator) specific" (277). SJT maintains, however, that this step is critical for the success of the analysis and warns that it is important to identify all major cues for the problem.

For five of the six approaches, it is sometimes the judgment or decision analyst rather than the judge or decision maker who defines the problem. DT is the sole exception. Its sole intended function is use in applied decision problems, and it conducts no basic empirical research on judgment or decision processes. (All its basic research is mathematical or logical in character.)

BDT, PDT, SJT, IIT, and AT all do conduct basic empirical research concerning human judgment or decision making. In such research the judgment or decision problem and its dimensions are defined by the researcher. The procedures for this step are grist for the philosophy of science mill, and no attempt will be made to describe them here. Suffice it to say that the procedures are even

more situation and investigator specific in basic research than in applied work.

Fact/knowledge distinction. In applications, a basic procedural question is "who makes what judgments?" The three approaches actively involved in applications, DT, BDT, and SJT, all seem to agree that (a) policy makers should make judgments about values or preferences and (b) experts should make judgments about facts or inferences about environmental attributes. (See, from DT, Keeney & Raiffa, 1976, pp. 377-390; from BDT, Gardiner & Ford, 1977; from SJT, Hammond, Mumpower, & Smith, 1977.)

The procedures by which particular policy makers and experts are selected to make judgments are not explicitly laid out by any of the approaches, however. In general, the policy makers who make the value judgments are those somehow identified as individuals or groups with a legitimate voice in the decision. Experts are usually identified on a more ad hoc basis, although they ordinarily possess recognized stature within their discipline.

Type of response. The procedural issue of concern here is the nature of the response demanded by the judgment or decision task. There is great variety both within and across approaches in the procedures used to elicit responses from judges or decision makers.

It is useful to first distinguish among the various foci of different types of tasks before turning to the specific modes of response required. First, five of the six approaches are frequently interested in judgments or decisions related to preference, value, or utility. These are DT, BDT, PDT, SJT, and IIT; AT ordinarily

confines its interests to the manner in which judges make decisions about states of the world (inferences). A further useful distinction in this regard is whether or not the procedures involve making a choice for action. In DT, the intended function of which is to aid decision makers to make decisions in a "rational" manner, tasks always require the decision maker eventually to make a choice for action among multiple alternatives. BDT and SJT, in their applied versions, are also intended to help decision makers make choices for action. BDT, PDT, SJT, and IIT, however, unlike DT, often construct tasks in which the aim is not aiding the decision maker and in which no choice for action is ever made. Rather, the aim is simply to describe, predict, or explain the manner in which the judge uses information in making a judgment of preference.

Second, five of the six approaches are frequently involved with inferences about states of affairs in the world. These are BDT, PDT, SJT, IIT, and AT. Such inferences may have implications for action or choice, but not necessarily so. In such tasks, the judge is asked to express his inference about the "true" state of the world, given certain pieces of information. For instance, the judge may be asked to estimate the value of a distal variable or to make an attribution concerning characteristics of an individual on the basis of a set of cue or attribute values.

Finally, and really a special case of the second type of task, three of the six approaches (DT, BDT, and PDT) quite frequently construct tasks in which the judge is required to make probability estimates. These may be either estimates of the probability of an

event occurring, estimates of the probability of a hypothesis given an event, or estimates of the probability of an event given an hypothesis. Such tasks appear both in basic research and in practical applications. DT, BDT, and PDT are ordinarily concerned with such "risky" decision making, or decision making under conditions of uncertainty, although DT and BDT are increasingly also concerned with "riskless" decision making, or decision making under conditions of certainty.

SJT, IIT, and AT are ordinarily concerned with riskless decision making, but exceptions can be cited for all three approaches. IIT has devoted substantial attention to risky decision making (e.g., Shanteau, 1975); SJT has occasionally constructed tasks requiring judges to give probability estimates (e.g., Hammond, Stewart, Adelman, & Wascoe, 1975); and, AT appears to be moving toward including tasks involving probability estimation (e.g., Ross, 1977). In terms of procedures, however, it is still more likely to find questions of the sort, "How likely is it that. . .?", in the first three decision or "probability" theories than in SJT, IIT, or AT.

We now turn to the specific procedures most frequently found in each of the three types of tasks: those involving judgments of preference or utility, those involving judgments of inference, and those involving probability estimation.

Judgments of preference or utility. Within DT, in tasks concerning judgments of preference the decision maker's responses can take several forms. Most often they are written or verbal responses

to questions about the decision maker's ordinal preferences between pairs of alternatives, particularly alternatives involving lotteries. For instance, decision makers are commonly asked to express their ordinal preferences between two lotteries of the form "receive outcome X with probability p or outcome X' with probability $1-p$," or between an option of the form "receive consequence C with certainty or receive a lottery offering outcome X with probability p or outcome X' with probability $1-p$." The purpose of these questions is to identify pairs to which the decision maker is indifferent. The responses to these questions can then be used in determining whether certain necessary theoretical axioms for DT are satisfied, and, if so, to construct a model of the judge's utility function. The specific procedures used to elicit these responses (and variations in the types of alternatives presented and responses required) are described in greater detail in later sections. The important point is that the basic tasks presented the decision maker requires only the expression of ordinal preferences between alternatives; the parameters of the model of the decision maker's utility function are then approximated from those responses. (Discussion of the nature of those parameters is deferred to later sections.)

The last point is important because BDT differs substantially from DT in the procedures it uses to estimate the parameters of the decision maker's utility function. Although DT and BDT share similar theoretical orientations, rather than deriving the parameters of the utility function indirectly from expressions of ordinal preference, BDT frequently requires the decision maker to specify directly

the value of those parameters. That is, the decision maker is asked to indicate directly the relative importance of the various attributes of the decision problem and to describe directly the form of the value or utility curves indicating the functional relation between values of each attribute and the decision maker's utility. This direct specification procedure is used in recent applications of the BDT approach; basic research by BDT frequently relies upon techniques more similar to those described above for the DT approach.

PDT relies extensively on the ordinal preference method described above for the DT approach. The judge is most commonly asked to make choices between two lotteries, between a certainty equivalent and a lottery, or between complex combinations of lotteries and/or certainty equivalents (e.g., Kahneman & Tversky, 1977). PDT has also at times used procedures involving wholistic ratings of utility or desirability, similar to those described below for SJT (e.g., Slovic & Lichtenstein, 1971).

Within SJT, in tasks involving judgments of preference or utility the judge is ordinarily required to make wholistic ratings of preference over a series of cases in which the values of each attribute vary. Such ratings are made on interval scales. They are sometimes recorded in written form, although more commonly the series of cases are presented and the judge's responses recorded via interactive computer graphics techniques. SJT is the only one of the six approaches to rely heavily upon computer graphics technology, although interactive computer programs have recently been developed within the DT approach (e.g., Leal & Pearl, 1977).

Occasionally, judges are requested to specify directly the values of the parameters of models of their judgment processes, in a manner similar to the procedures of BDT. Judges may be asked, for example, to indicate the relative importance of cues by dividing 100 points among them or to indicate the relation between levels of an attribute and desirability by drawing graphs.

The IIT approach relies almost entirely upon wholistic ratings of specific cases in its study of judgments involving preference. The judge responds on what is assumed to be an equal interval scale (that assumption is routinely tested in IIT). Responses are ordinarily recorded by the judge in a paper and pencil format.

Judgments of inference. All the approaches, with the exception of DT, at times construct tasks in which the judge is required to make inferences about unobservable, distal variables on the basis of observable, proximal information. (DT does occasionally require experts to make subjective evaluations of the values of attributes on scales such as, say, "quality," but such tasks appear infrequently and are not an integral part of the approach, therefore, their omission here.)

Tasks from the BDT, PDT, and AT approaches often require the judge to respond to paired choices. For example, in the BDT and PDT approaches, judges are commonly asked questions of the form, "Given datum x, is event y or event z more likely to occur?" AT tasks are frequently of the form, "Given event x, is cause y or cause z more likely?"

PDT, SJT, and IIT frequently construct tasks which require judges to express their inferences in terms of evaluations or ratings. For example, tasks within these approaches frequently ask judges to express their judgment about the most likely value of some unobservable object on the basis of a set of given cue values or attribute levels. IIT and AT also frequently construct tasks in which the judge is presented a number of attributes describing an individual and is required to evaluate that individual on another attribute dimension (i.e., tasks of impression formation).

Probability estimation. Probability estimation tasks, and the responses required by them, can be roughly subdivided into two categories: those requiring indirect responses and those requiring more or less direct estimation of probability or likelihood. The indirect procedure requires the decision maker to express his ordinal preference for alternative lotteries. Subjective probability estimates of events are then derived from these responses. This method is used extensively within the DT approach. It is also found frequently in basic research within BDT, although it appears infrequently, if ever, in the practical applications of BDT.

The second type of task requires the judge or decision maker to express his subjective estimates of probability more directly. This type of task is used by all three of the decision approaches, DT, BDT, and PDT. It is also the approach used by IIT in their studies of risky decision making and by SJT in its infrequent studies involving probability estimation. These tasks may require the decision maker to make a statement in terms of numbers between

0 and 1, or to express the odds for various events (e.g., "how much more likely is it that x_1 will occur instead of x_2 ? Twice as likely? Three times?", etc.), or to construct (with the aid of the decision analyst) a probability distribution. This last procedure is usually based on the fractile method. The decision maker is first asked a series of questions designed to specify the value of attribute X , X' , such that he feels it is equally likely that the obtained value of X will be between X_0 (lowest feasible value) and X' or between X' and X_1 (highest feasible value). This procedure is reiterated for the intervals (X_0, X') and (X', X_1) and so forth, until enough points have been plotted to sketch the distribution.

Design. The design of studies of judgment and decision making was covered extensively in the previous section. Design has direct implications, however, for procedures in so far as it affects the nature of the materials the judge or decision maker sees and the manner in which he is required to respond. The most important design factors affecting procedures are related to the idiographic versus nomothetic distinction discussed above in the Methodology section.

For five of the six approaches, DT, BDT, PDT, SJT, and IIT, their aims are sometimes, if not always, idiographic in orientation. This orientation requires that sufficient data be collected from each judge or decision maker in order to construct a model of that individual's judgment or decision processes (a model which may or may not be imputed paramorphic qualities). This usually, but not always, entails factorial design in the within-subjects component

of the study or analysis. That is, each judge is required to respond to every possible combination of stimulus attributes or to all possible combinations of a finite number of levels of the stimulus attributes. In terms of procedures, this means the judge or decision maker is required to make repeated judgments or decisions.

The aim of DT is exclusively idiographic. Procedurally, the decision analyst will ordinarily construct tasks containing large numbers of attribute level combinations to which the decision maker will theoretically be asked to respond. The approach can thus be described as factorial, although it is not referred to or described as such within DT. The DT approach usually involves "short cuts," however. For example, instead of asking the decision maker to respond individually to each of a large number of stimulus attribute combinations, say, to indicate his preference between attribute combinations (X,Y) and (X',Y') , over a whole series of values, Z_1, \dots, Z_n , the decision maker may be asked to indicate his preference for (X,Y) versus (X',Y') for Z_1, Z_2, Z_k and Z_n , and, as soon as he understands the nature of the problem, be asked directly if his preference for (X,Y) versus (X',Y') would change as a function of any plausible value of Z . DT procedures are dynamic, in the sense that the choices presented by the analyst to the decision maker are derived from responses to earlier choices.

Although BDT, in its practical application function is idiographic in orientation, it does not use a within-subjects factorial approach to collect data for the construction of a model representing

the decision maker's decision processes. Rather, the analyst asks the decision maker to indicate directly the characteristics of the model, unlike the procedures of all other approaches which involve responses by the judge from which such models are indirectly derived. For instance, decision makers are asked to specify the relative importance of attributes by assigning weights on a ratio scale.

In the basic research component of BDT, and in PDT, judges respond to each of a pre-established array of stimulus combinations, although that array may not be the same for every judge. Judges may or may not see every possible combination of levels of stimulus attributes, but more often they do not. Each combination of stimulus attributes, however, is seen by some judges. The primary orientation of basic research in these two approaches is more nomothetic than idiographic. The design of basic research in these two approaches, then, is usually between-subjects factorial. The researcher establishes a set of task materials that include all possible combinations of a number of levels of attributes. Each judge responds only to a set or subset of those combinations; judges often respond to only one instance of any particular judgment or decision problem. Although models may be tested for individual judges, the tests of interest is ordinarily between-subjects.

In both the applied and basic versions of SJT, the basic orientation is always idiographic. Again, this means that individual models are constructed for each judge and that judges respond to multiple instances of a particular judgment problem. Unlike

most other approaches, however, construction of individual models does not involve factorial design. Because SJT relies extensively on regression techniques in model building, the judge ordinarily responds to only a sample of all possible combinations of cue levels. In terms of procedures, the analyst or researcher first generates a sample of cue level combinations. The judge then responds to each of the cue level combinations (or profiles) in this pre-established set. Unlike the factorial approaches, in which the stimulus attributes are almost always orthogonal, this is frequently not the case in SJT. The SJT approach emphasizes the concept of representative design. That is, the cue level combinations to which the judge responds should be representative of those in the "real world," preserving cue intercorrelations, ranges and distributions, and other characteristics of the cues in the natural environment. Such characteristics are frequently preserved in the generation of the set of cue level combinations. Basic research within SJT frequently makes use of between-subjects factorial designs, but only after within-subjects analyses have been conducted.

IIT relies exclusively on factorial design, in which each judge or some group of judges see each combination of stimulus attribute combinations. In order to keep the total number of cases to a manageable level, either the number of attributes or the number of levels of each attribute is ordinarily small in the typical IIT study. Because of the reliance on factorial design, attributes are ordinarily orthogonal within sets. IIT uses both within-subjects

and between-subjects factorial designs. Most often, within-subjects analysis precedes between-subjects analysis.

AT is the only one of the six approaches that rarely, if ever, constructs models of individual judgment or decision processes. Its primary aim is nomothetic. It therefore uses between-subject factorial designs extensively. In terms of procedures, it first constructs a set of attribute level combinations, which are almost always orthogonal. Most commonly, each judge responds to only one of these combinations (or to only a small subset). It is quite common in AT, since the approach does not require extensive data from individual judges in order to build models of their judgment or decision processes, for judges to respond to a number of different judgment and decision problems within a single study.

[MORE]

Level of abstraction. After the nature of the judgment or decision problem has been established and the structure and dimensions of the task determined, a basic issue is the level of abstraction at which the attributes or cues will be presented to the judge. This issue was discussed at length in the previous section on methodology, but a brief discussion is also appropriate here.

The level of abstraction at which attributes are presented ranges on a continuum from concrete, realistic objects to abstracted variables. The approaches using the most concrete (realistic) level of presentation are PDT, IIT (sometimes), and AT. All three approaches sometimes use very detailed verbal descriptions of objects. IIT has used photographs, as has AT, which at times has

even used recordings or videotapes. The next level of abstraction is that of partially decomposed objects, i.e., real or hypothetical objects, presented in schematic terms. DT, BDT in its research function, SJT, and IIT (usually) fall into this category. BDT, in its applied function, is the only approach to operate procedurally with abstracted variables or dimensions themselves.

The greatest implication for procedures of the level of abstraction of variables is the number of cases that are presented for consideration by individual judges. More abstract level of presentation enables stimulus combinations to be manipulated more easily and require less time for presentation and response. As noted in the methodology section, within-subjects designs are therefore usually characterized by more abstract levels of presentation than are between-subjects designs.

Cue metric. An issue closely related to that of level of abstraction is the metric in which the attribute levels are presented. In general, the more abstract the level of presentation, the more likely that the levels will be represented in numerical form. DT, BDT, SJT, and IIT, therefore, more commonly represent levels of attributes by numbers than do PDT and AT.

SJT almost always presents attributes in quantitative form, with an underlying continuous distribution. Such a metric is dictated by the requirements of the regression analysis techniques ordinarily used by SJT. Techniques for application of SJT in non-metric (i.e., nonnumeric) ecologies have been developed (Bjorkman, 1973; Castellan, 1970), but are not extensively used. DT, BDT,

PDT, and IIT all at times present attributes in similar quantitative form; AT rarely, if ever, does so. All of the approaches sometimes present stimulus attributes in discrete forms or in terms of qualitative descriptions, although SJT does so less commonly than the rest.

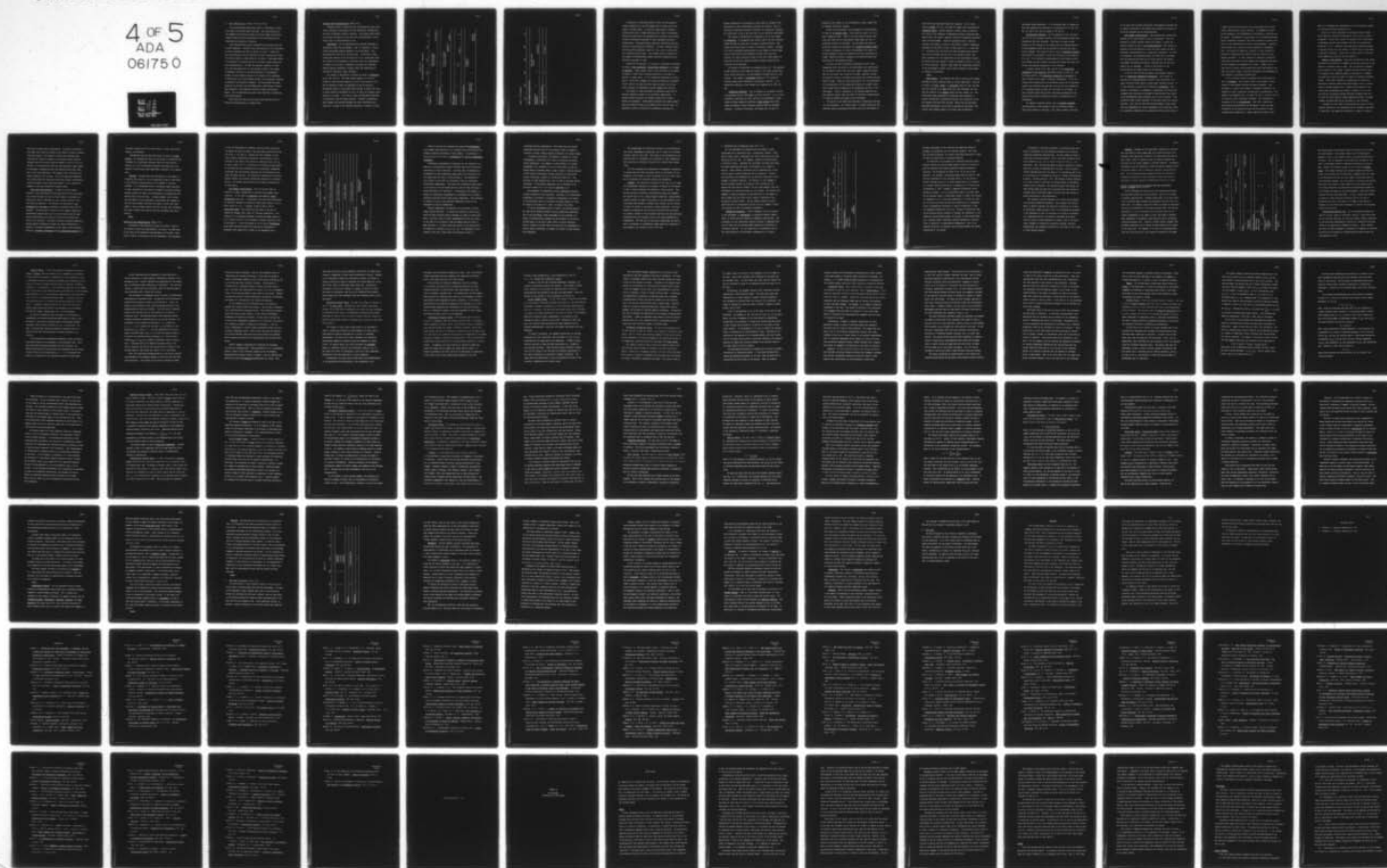
Overview. The six approaches show considerable diversity in the manner in which they construct tasks and present stimulus attribute or cue combinations to judges and decision makers. They also vary greatly in the types of responses they require. Several factors appear to contribute to these differences. One obvious contribution to differences among the approaches is differences in aim or function. The types of procedures employed are affected according to whether basic research or practical application is the intended function and according to whether the aim is more idiographic or more nomothetic. In addition, procedures vary as a function of whether the analysis is concerned with judgments of preference, judgments of inference, or probability estimation. Finally, differences in origins and differences in means of data reduction or data analysis probably account for some of the differences among approaches. The type of data most amenable to analysis by Bayes theorem, regression techniques, and ANOVA techniques vary, and thus the procedures involved in data collection also vary. All six approaches, however, construct tasks which provide data which can be used to construct models of the judgment and decision processes of individuals or of groups of individuals.

AD-A061 750

COLORADO UNIV BOULDER CENTER FOR RESEARCH ON JUDGME--ETC F/G 5/10
THE COLORADO REPORT ON THE INTEGRATION OF APPROACHES TO JUDGMEN--ETC(U)
OCT 78 K R HAMMOND, G H MCCLELLAND N00014-77-C-0336
CRJP-213 NL

UNCLASSIFIED

4 OF 5
ADA
06/75 0



B. Task Characteristics (Tables IV-B1 and IV-B2)

All six approaches assume that there is "something out there" with which the decision maker interacts. The characteristics of the judgment and decision problems with which judges and decision makers interact and the implications of such characteristics for procedures are therefore of interest.

Task characteristics can be subdivided into objective and subjective categories. Objective task characteristics can be described as those concerned with "what is really out there." Subjective task characteristics can be described as those concerned with "what the decision maker or judge thinks is out there." Approaches differ in the amount of attention they pay to objective versus subjective task characteristics. SJT pays the greatest attention to objective task characteristics, in part perhaps because of the legacy of Brunswik's interest in the characteristics of environmental systems, in part perhaps because of its origins in the psychology of perception, and in part perhaps because of the approach's interest in multiple cue probability learning. IIT pays perhaps the least attention to objective task characteristics, restricting its interests primarily to the cognitive systems of judges and decision makers. All six approaches pay quite close attention to subjective task characteristics.

Since objective task characteristics seem logically prior to subjective characteristics, we begin there.

Objective Task Characteristics (Table IV-B1)

Concepts related to objective task characteristics have little direct influence on procedures for most approaches, although they may have considerable indirect influence. We consider them briefly and discuss the manner in which they are operationalized, to the extent they are operationalized at all.

Uncertainty. All six approaches have concepts referring to uncertainty within the environment. That is, ambiguity or probabilism intervening between elements of the environment. This uncertainty may take one of several forms. It may entail uncertainty about whether or not an event will occur, uncertainty about the validity of an hypothesis, uncertainty about the consequences or effects of an action, or ambiguity concerning the relations between directly observable, surface events and unobservable or nondirectly observable, depth variables.

The concept of uncertainty is defined in terms of probability by DT, BDT, and PDT. (The same concept appears in IIT in its studies of risky decision making, in a few of the studies of SJT, and is beginning to make its appearance in the AT approach.) Probability refers to uncertainty about whether an event will occur, or about whether an hypothesis is true, or about the probable consequences of an event. It is the so-called structuralist definition of probability (see the Theory section). Because of probability, many judgment and decision problems are risky; uncertainty surrounds the outcomes of the possible decisions or courses of action.

Table IV-B1
Task Characteristics - Objective

| DT | BDT | PDT | SJT | IIT | AT |
|--------------------------------|----------------------------|-----|---|---------|-----------------------|
| Probability | X | X | Task Uncertainty (surface/depth) | | Covariation |
| <u>Organizing principles</u> | | | | | |
| Bayes theorem | X | X | Additive, weighted averaging models (configural model) (schema) | | |
| <u>Task elements</u> | | | | | |
| datum/events | X | X | Cues | Stimuli | X X |
| <u>Cue-criterion relations</u> | | | | | |
| X | Conditional probability | X | Ecological validity | | Multiple causality |

Table IV-B1 continued

| DT | BDT | PDT | SJT | IIT | AT |
|------------------------------|-----|-----|----------------------------|------------|----|
| Task elements interrelations | | | | | |
| Independence/ dependence | X | X | Cue intercorre- lations | Redundancy | |
| | | | | | |
| Number of task elements | | | | | |
| X | X | X | | Set size | |

Probability is important within DT, BDT, and PDT primarily because probabilism in the environment must be coped with by the judge or decision maker. In practical applications, objective probabilities have no common implications in terms of procedures. Objective probabilities are unknowable with certainty and interest focuses on subjective estimates of those unknowable objective probabilities. Procedures center around the elicitation and calibration of those subjective estimates. For basic research within the BDT and PDT approaches, tasks are constructed so that objective probabilities are known. The eventual issue of interest is the degree of correspondence between objective probabilities and subjective estimates of them.

Within the SJT approach, the concept of uncertainty is defined in terms of task uncertainty, surface/depth relations, or the zone of ambiguity. These concepts refer to the inability of the judges to predict or know distal variables perfectly on the basis of surface, or given, data. It is a frequentist definition of probability. Within SJT, judgment and decision problems are considered "risky" only in the sense that the judge can never be sure that there is perfect correspondence between judgment and criterion.

The concept of task uncertainty is important within the SJT approach primarily because of its behavioral implications. In applied decision problems, the concept usually has no direct effects on procedures. Surface/depth relations are usually unknowable and interest focuses on the judge's use of surface cues. In laboratory studies of multiple cue probability learning, the

primary implication for procedures is that tasks are (almost) never constructed so that surface/depth relations are certain. That is, the tasks in laboratory studies of multiple cue probability learning are constructed in such a way that they are not perfectly predictable.

The concept of uncertainty is recognized by AT in the concept of covariation. AT recognizes that similar events can have different causes and that the same cause can lead to multiple effects. The concept has little impact in terms of procedures, however. Although AT is quite interested theoretically with the manner in which judges come to learn about the environmental systems with which they interact, AT rarely constructs tasks in which the "right answer" is known and in which the relations between stimulus attributes and criterion can be specified.

No concept of uncertainty is prominent within IIT. This approach focuses primarily on the cognitive system of the judge, paying relatively little attention to the environmental systems with which they interact. The concept of probability appears in IIT studies of decision making under conditions of uncertainty, however. It is essentially identical to the concept as it appears in DT, BDT, and PDT.

Organizing principles. Also of interest is the manner in which the approaches see the environment as being organized. For the three probability approaches, DT, BDT, and PDT, the environment is often viewed as being organized according to Bayes theorem, which describes the manner in which objective probabilities are optimally combined. In terms of procedures, this means that in tests of

optimality with respect to the environmental system, judges will be compared with Bayes' theorem.

Within the SJT approach, the environment is often described in terms of an additive model. (By an additive model, we do not mean a summative model. That is, the term additive merely indicates that there are no cross-product terms among the cues or attributes involved in the model. The SJT model could also be described, perhaps more precisely, as a weighted averaging model.) In terms of procedures, in tests of achievement with respect to the environmental system, the additive model best describing the judges' cognitive system is compared to the additive model best describing the environmental system.

For both the three probability approaches and SJT, their assumptions about the nature of the environment have implications for the manner in which they deal with judges' cognitive systems. That is, the models they construct of judges' cognitive systems are similar in form to the models they assume best describe environmental systems. Implicit in the procedures of all four approaches is the assumption that the organizing principle of judges' cognitive systems should correspond to the organizing principle of the environment. Tests of congruence or fit between judges and the environmental system, then, are tests of similarity between models of cognitive systems and models of the environment.

IIT and AT do not explicitly postulate an organizing principle for the environment. By "reverse logic," we might assume that they postulate similar organizing principles within the environment as

those which they postulate within the organism. For AT, these would be schema; for IIT, they might be (among other possibilities) configural models. Neither approach, however, seems to maintain that there is any necessary correspondence between organizing principles within the judge's cognitive system and organizing principles within the environment. These two approaches seem to take the position that judges, at least to some extent, "impose" organizing principles on the world in order to make sense of it, or cope with it. (We should note that while the above statements hold more-or-less true for present day AT, one of its founders, Heider, was quite concerned with the organization of the environmental system. His views were quite similar to those of Brunswik and thus, to those of the SJT approach. Heider's views about the organizing principles of the environment, however, seem to have relatively little impact on current AT procedures.)

[MORE]

Task elements. The elements that make up decision and judgment problems are given different names by various approaches. They are referred to as data, events, or attributes within DT, BDT, and PDT; they are referred to as cues within SJT, and, sometimes, AT; and, they are referred to as stimuli within IIT and, usually, AT. The differences among approaches in terminology seem primarily to reflect differences in theoretical origins, and conceptions of how the organism interacts with the task. Aside from the previously discussed differences in the level of abstraction and metric, the various task elements appear to be pretty much procedurally

equivalent across approaches. It is relatively easy to imagine the same task elements being referred to as attributes by DT, BDT, and PDT, as cues by SJT, and as stimuli by IIT and AT.

Cue-criterion relations. All the approaches, with the exception of IIT, have concepts related to the relationship between task elements and the task criterion. This type of concept is most prominent in SJT because of its concern with the characteristics of the environmental system and its interest in multiple cue probability learning. The concept in SJT is that of ecological validity, which refers to the degree to which the criterion can be predicted by a cue. It is ordinarily operationalized as the simple correlation between cue and criterion (although the concept does take into account the possible nonlinearity of the relation).

A similar concept in DT, BDT, and PDT is the conditional probability of the criterion (hypothesis) given an event (or, some attribute level) or the conditional probability of an event (or, some attribute level) given a criterion (hypothesis). In DT, the concept has no implications for procedures because of its exclusive concern with practical applications, in which the true probability of a hypothesis given an event (or vice versa) is not known. In basic research within BDT and PDT, conditional probabilities concerning cue-criterion relations are known or manipulated by the experimenter.

AT contains a related concept, that of multiple causality. The implication of this concept is that the relation between a cause and an effect is uncertain. But, aside, perhaps, from some

of the early work in person perception, environmental relations are not known or specified in AT research and concepts concerning cue-criterion relations are not operationalized.

Task element interrelations. The interrelations between task elements are of some interest to all six approaches. They are frequently of great interest within SJT, however, in which the relevant concept is that of cue intercorrelation. This concept is defined as the predictability of one cue from a second cue and is ordinarily operationalized as a correlation coefficient. In terms of procedures, cue intercorrelations are often set in accordance with the magnitude and direction of those found in the natural environment (representative design), although they are also frequently set as orthogonal. Studies of the effects of cue-interrelations are common within the SJT approach.

In the DT, BDT, and PDT approaches, the relevant concept is that of conditional dependence/independence. This concept is defined as and operationalized in terms of the probability of an event (attribute level) given a second event (attribute level). It is also referred to within the DT approach as redundancy. With respect to procedures, in judgments of preference, both DT and BDT recommend that attributes be defined or constructed so that conditional dependencies among attributes are as low as possible. This procedure avoids "double counting" of attributes within the utility function. Also, for DT, such task structure makes it more likely that the independence of utility assumption can be satisfied, which is a necessary condition for the construction for each attribute of

a single utility function which can be incorporated into a multi-linear, multiattribute utility function. In judgments of probability estimation, the independence of attributes is ordinarily not of concern. The emphasis in application is on calibrating decision makers' probability estimates, and little attention is given to the nature of the environmental sources of those estimates. Moreover, Bayes theorem can be used with either dependent or independent data. For reasons of simplicity, however, in basic research on probability estimation, BDT usually constructs tasks with independent attributes. In basic research on probability estimation within PDT, both tasks in which elements are independent and tasks in which elements are dependent can be found. Indeed, much of their work is focused toward demonstrating that individuals do not take into proper account conditional dependence and independence of task elements in estimating probabilities.

In IIT, the major concept concerning task interrelations is that of redundancy. This concept refers to the degree to which stimuli or task elements provide the same information. Such redundancy is usually self-evident or "apparent" redundancy, not measured in terms of empirical correlations or probabilities. For example, the statements "Mr. A is kind" and "Mr. A is kind to Mr. B" are obviously (here, logically) redundant. The counterpart to redundancy in IIT is inconsistency. Wyer (1970) defined and operationalized inconsistency as the degree to which the joint probability of two stimuli are less than the product of their unconditional probabilities. Several empirical studies on the

effects of redundancy and inconsistency on the information integration process have been performed by IIT.

There is no clear counterpart to the above concepts within the AT approach, although the importance of this idea is clearly recognized. Kelley (1971) noted, for instance, that "conflicting or ambiguous cues may evoke schemata that lead to contradictory inferences" (p. 173). Moreover, there has been a substantial amount of research within AT on the effects of "mixed" sets of adjectives in impression formation. (This work is quite similar to that described immediately above for the IIT approach.)

Number of task elements. There has been empirical work across approaches on the effects on judgment and decision processes of the number of attributes within a task. The IIT approach is perhaps most prominent in this regard. Numerous studies involving tests of averaging versus summation models as descriptors of judgment processes have varied the number of items of information presented judges. Within BDT's research on conservatism in probability estimation, numerous studies have varied sample size as a simple means of varying diagnosticity. Within the PDT approach, some studies have investigated the effects of number of cues on cue utilization; the general finding (although a somewhat weak one) was that the predictability (using regression methods) of judges' judgment processes declines as the number of cues increases.

Although it is not clear to what extent the results of basic research motivate it, all approaches seem to recognize, explicitly or implicitly, the cognitive limitations of judges in trying to

deal with too many items of information. In terms of procedures, this means that tasks are limited in the number of items of information they contain. For instance, DT explicitly advocates that attributes be limited in number to the greatest extent possible, analyses within SJT usually limit the number of cues to 5-8 or so, and BDT, in its application function, also advocates a small number of attributes, although setting a more generous upper limit than most of the other approaches. BDT suggests that 8 attributes are enough, whereas 15 are too many. If all the important dimensions of a problem cannot be contained within these limits, the generally accepted procedure across approaches appears to be to expand the judgment or decision problem into multiple tasks.

Other task characteristics. A number of other concepts related to task characteristics have been omitted from this discussion, either because they have no direct impacts on procedures or because they affect procedures for only a small fraction of the work within an approach. For instance, the importance of cue variability and cue range is explicitly recognized within the SJT approach. SJT warns that the results of any judgment analysis are theoretically generalizable only to cue sets exhibiting the same formal characteristics as those of the cue sets with which the analysis was conducted. (DT offers a similar procedural admonition, warning against attempting to assess the relative importance of weights of attributes independently of the range of the attributes.) Similarly, ecological reliability and cue intersubstitutability are

important concepts in SJT, but they rarely, if ever, have direct impacts on procedures.

Considerable work has been done within IIT and AT on order effects. But although the order in which items of information are presented is an important procedural feature of many specific studies, it is not an important procedural consideration in the majority of work within these approaches; therefore, it is omitted here.

Overview. Although they vary considerably in the degree of attention they accord it, all six approaches attend to some extent to the objective characteristics of the judgment or decision problems. To a considerable degree, differences among approaches in their conception of the nature of judgment or decision problems probably account for some of the differences in procedures we will discuss later in this section. A strong argument could be made that the manner in which approaches conceptualize the judgment or decision task strongly influences the manner in which they think judges or decision makers should cope with the task and, ultimately, the types of models they construct and the procedures they use in doing so.

[MORE]

Subjective Task Characteristics (Table IV-B2)

In addition to the question of "what's out there," there is the question of what the judge thinks is out there. We refer here to the issue of the subjective characteristics of the task. This issue is really at the heart of all six approaches. The procedures

of all the approaches are ultimately directed toward clarifying the manner in which the judge or decision maker perceives the task and tries to cope with it. Because the full treatment of this issue requires considerable elaboration, best deferred to later sections of the report, this section is relatively brief, as can be seen in Table IV-B2. In particular, the subjective equivalents to the objective task characteristics of uncertainty, organizing principles, and cue-criterion relations are all discussed in the three following sections. The subjective counterparts of the task element concept have received considerable attention in the Theory section. We discuss only subjective task element interrelations in this section.

Task element interrelations. Both SJT and PDT share an interest in judges' perceptions of objective task element interrelations. In SJT these are perceived (environmental) inter-correlations among cues. Procedurally, perceived intercorrelations are generally assessed by having the judge predict levels of cue X, given levels of cue Y. A similar concept in PDT is that of associative bonds. This concept has its basis in the work of Chapman and Chapman (1967, 1969) on "illusory correlation." Procedurally, it is generally assessed by having the judge predict or try to recall the level of an attribute X, given a particular level of attribute Y. (A related concept in IIT is that of discounting, which holds that when stimuli are perceived as correlated or dependent, less weight will be given to the aggregated pair.)

Table IV-B2
Task Characteristics - Subjective

| DT | BDT | PDT | SJT | IIT | AT |
|--|-----|----------------------|----------------------------------|-------------|----|
| <u>Uncertainty</u> | | | | | |
| See Theory section, following sections | | | | | |
| <u>Organizing principles</u> | | | | | |
| See following sections | | | | | |
| <u>Task elements</u> | | | | | |
| See Theory section | | | | | |
| <u>Cue-criterion relations</u> | | | | | |
| See following sections | | | | | |
| <u>Task element interrelations</u> | | | | | |
| Independence/ dependence | X | Associative bonds | Perceived inter- correlations | Discounting | |
| | | X | | | |

While SJT and PDT are concerned with perceived environmental task element interrelations, DT is concerned with the interrelations between subjective transforms of task elements. In particular, a central issue in DT is that of preferential and utility independence/dependence.

Preferential independence is necessary for the creation of an additive or multiplicative utility function (see later sections for discussion of utility functions). The basic idea of preferential independence is that preferences among levels of an attribute X do not depend upon the level of some attribute Y. The extension of this concept to more than two attributes is that of conditional preferential independence. This means that preferences among levels of an attribute X do not depend upon the level of some attribute Y, given some level of an attribute Z. Procedurally, the presence or absence of preferential independence is established via a series of questions involving choices among paired comparisons. This procedure is similar in form to that described immediately below for the assessment of utility independence.

Utility independence is necessary for the creation of multi-linear utility functions. It allows the creation of utility functions for each attribute without reference to levels of other attributes. The basic idea of utility independence is that the utility of any level of an attribute X does not depend upon the level of some attribute Y. In other words, the utility of any level x does not change as a function of y, y', y'' , etc. The extension of this concept to the case of more than two attributes is that of

conditional utility independence. This means that the utility associated with any level of an attribute X does not change as attribute Y varies, holding constant attribute Z at various levels.

In terms of procedures, the presence or absence of utility independence is established by a series of questions involving paired comparisons. For example, to check whether X is utility independent of Y, attribute Y is set at some level, y , and the analyst asks the decision maker to make a series of choices between a 50-50 lottery (involving x values) and either a second 50-50 lottery (involving x values) or a particular level of X. The procedure is then repeated with Y set at level y' and then further reiterated. If the paired comparisons are not affected by the particular levels of Y, X is utility independent of Y.

The assessment of conditional utility independence involves a simple extension of this procedure. Attribute X is conditionally independent of attribute Y (although not necessarily vice versa), if the preference order for lotteries involving changes in X does not depend on the level of Y, when Z is fixed at any level.

If utility independence does not obtain, the decision analyst may take any of several courses of action, among them transformation of the attributes, direct assessment of the utilities of a number of multiattribute outcomes and extrapolation to the entire outcome space, or the use of other forms of utility functions. Since it appears to be most often the case that the independence of utility axiom is satisfied, we comment no further on these alternative techniques.

BDT acknowledges the theoretical validity of the preferential and utility independence assumptions, but, at least in its applied function, does not test them. BDT relies on the robustness of the additive model to accommodate any violations of these assumptions. There are thus no procedures within BDT for testing these assumptions.

IIT also recognizes the importance of independence assumptions. It assumes that the scale and weight values of attributes are not dependent on the levels of other attributes, but, as in BDT, there are no procedures for routinely testing this assumption.

Overview. The subjective properties of the task for the judge have considerable influence on procedures, although the discussion of most of these influences was deferred to later sections. This section focused on subjective interrelations among task elements. SJT and PDT (and IIT to a lesser degree) define and operationalize concepts related to judges' perceptions of the environmental relations among task elements, although these concepts are not central to either the theory or procedures of these approaches. The independence of preferences and utilities associated with attributes is, however, central to the DT approach and there are well developed procedures which are routinely used to test these assumptions of independence. Both BDT and IIT explicitly acknowledge the importance of assumptions of independence among subjective transforms of task elements, but neither routinely test them.

C. Subjective Use of Subjective Data (Table IV-C)

All six approaches are concerned with the manner in which judges make use of subjective data. Procedurally, however, "soft" data is almost always transformed into "hard" data before the judge interacts with the task. For example, consider the specification of attribute levels. Sometimes the environment provides hard objective data concerning the levels of attributes, e.g., number of dollars. Often, however, attributes can be expressed only in subjective terms, as with attributes such as "esthetics." Usually, such attributes are translated into an objective form prior to the judgment or decision analysis. For instance, if the judgment or decision problem includes an attribute such as esthetics, the analyst may first have an "expert" (or the judge himself) rate each alternative on a one-to-ten scale of, say, beauty. The scope of the judgment or decision analysis is then restricted to the manner in which judges deal with these "hard," numeric transforms of the subjective attribute beauty. The one common exception in which subjective data is incorporated directly into the judgment or decision analysis is that of probability estimation.

Probability estimation. A major component of the DT approach is the refinement, or calibration, of subjective hunches, impressions, or intuitions into subjective judgmental probabilities. The procedures for such calibration involve lotteries or imaginary bets and are based upon the decision maker adopting two principles of consistent behavior: (a) the transitivity of preferences and (b) the substitutability of indifferent consequences in a lottery.

Table IV-C
Subjective Use of Subjective Data

| DT | BDT | PDT | SJT | IIT | AT |
|---------------------------------------|----------------------|-----|-----|-----|----|
| <u>Attribute level transformation</u> | | | | | |
| X | X | X | X | X | X |
| <u>Probability estimation</u> | | | | | |
| Lotteries | | X | | | |
| Calibration | | | | | |
| | direct estimation | X | | | |

Internal consistency is the criterion for admitting subjective estimates of probability into the decision analysis. PDT often uses the same procedures, but without requiring individuals to adopt the above two principles of consistent behavior.

In particular, the procedures for assessing subjective probabilities involve the decision maker making a series of choices in which he indicates his ordinal preference between alternative lotteries. The procedures are based on an "if you had to bet" approach. For instance, the decision maker might be asked to indicate his preference for the following choice: "Do you prefer a lottery offering prize W if it rains tomorrow and L if it doesn't, or a lottery offering prize W with a probability of .20 and L with a probability of .80?" Formally, judgmental probability can be defined as follows: If E is an uncertain event in the real world and the decision maker is indifferent between a lottery that results in consequence X if E occurs and consequence Y if E does not occur and a lottery that gives an objective probability p of consequence X and a complementary probability of $1-p$ of consequence Y , then the judgmental probability of E is p . During the calibration task, the DT decision analyst attempts to identify and compensate for any cognitive or motivational biases on the part of the decision maker, e.g., personal biases, anchoring, unstated assumptions, conservatism, or inability to conceptualize independence. Much of the research within PDT is concerned with the specification and further description of such biases.

In addition to the above procedures for eliciting point estimates of probability, DT has developed procedures for eliciting subjective probability distributions. The most common of these is based upon the fractile method. First, the lowest plausible value of an attribute X , \underline{X}_0 , and the highest plausible value, \underline{X}_1 , are identified. Next, the decision maker is asked to express his preferences for a series of alternatives of the form: "Do you prefer a lottery offering prize \underline{W} if the value of X is between \underline{X}_0 and \underline{X}' and \underline{L} if the value of X is between \underline{X}' and \underline{X}_1 or a lottery offering prize \underline{W} with $p = .50$ and \underline{L} with $p = .50$?" The purpose of these questions is to find a value X' such that the decision maker is indifferent between the two lotteries. The intervals $(\underline{X}_0, \underline{X}')$ and $(\underline{X}', \underline{X}_1)$ are then divided using the same type of procedure and this process is iterated until a sufficient number of points have been plotted to permit description of the distribution.

BDT ordinarily and PDT sometimes rely on more direct methods for eliciting probability estimates. Most commonly the decision maker is asked to estimate the subjective probability of an event by specifying a number between 0 and 1, by indicating his estimate of the likelihood or odds (or log odds) of an event's occurrence, or by indicating which of two events he considers more likely. (SJT has relied on point estimation in numeric form in its infrequent treatment of probability estimation problems; objective probabilities are ordinarily specified for the judge in IIT's study of risky decision making.)

Overview. Although all six approaches' theories are concerned with the manner in which judges make use of subjective data, procedurally, most approaches "objectify" the subjective data used in their tasks. That is, subjective data is usually converted into hard data for the formal judgment or decision analysis. Probability estimation is the one exception in which procedures commonly involve the decision maker's use of subjective, personal, and nonexplicit data. Of the three approaches commonly dealing with problems of probability estimation, DT more frequently relies upon indirect procedures to elicit subjective probabilities than do DT or BDT.

D./E./F. Subjective Use of Objective Data/Cue Utilization/
Central Processes (Table IV/D/E/F)

We are combining our discussion of the procedures associated with the above three loci into one section. We do so as a result of several considerations. Foremost among these, it is quite difficult to associate many concepts and procedures with only one of the three loci. For example, should the concept of weight be associated with the cue utilization or the central process loci? The relative importance of a cue or attribute (cue utilization loci) is not totally independent of the manner in which the judge or decision maker aggregates or integrates across multiple cues or attributes (central processes loci). Moreover, closely related concepts from different approaches are not always located similarly with respect to the three loci. For example, if it were to be associated with only one of the three loci, the concept of valuation in IIT might

Table IV - D¹E/F

| DT | BDT | PDT | SJT | IIT | AT |
|------------------------------------|---------------------------|-----|-------------------|-------------------|------------------|
| <u>Utility/value/function form</u> | | | | | |
| value/utility | X | X | | | |
| | | | function form | valuation | (directionality) |
| <u>Weight</u> | | | | | |
| X | X | X | weight | X | (significance) |
| subjective probability | X | | | | |
| <u>Organizing principle</u> | | | | | |
| Bayes Theorem | X | X | | | |
| Additive model | X | | X | X | |
| X | Weighted averaging models | | X | X | |
| Multiplicative models | X | X | | X | |
| Multilinear models | | | X | | |
| | | | | | schemata |
| <u>Psychological factors</u> | | | | | |
| Biases | X | X | X | X | X |
| | | | cognitive control | (goodness of fit) | |

most appropriately be discussed under the subjective use of objective data heading. This concept refers to the "psychological meaning" of data, or the manner in which the judge rescales objective data. The SJT approach, on the other hand, includes no concepts or procedures related to subjective rescaling of objective data. Rather, the relevant concept in SJT is that of function forms. This concept describes the relation between the objective values of attributes and the individual's judgments and, if it were to be associated with only one loci, might be discussed under the cue utilization heading. But the concepts of valuation and function forms are quite similar in many respects (see the Theory section for further discussion of similarities and differences between the two concepts) and should be considered together, along with related concepts from the other approaches. In short, rather than attempt to impose an overly artificial structure and run the risk of distorting the facts of the matter, we judged it preferable to collapse into one section the discussion of concepts and procedures across the three loci.

Utility/value/function form. All six approaches possess concepts of utility, value, or functional relation. These three terms are not strictly identical, but they are formally similar, referring to the functional relationship of objective or subjective attribute values to the individual's judgment or decision. These concepts may refer to either judgments of preference or judgments of inference. We discuss the procedures for operationalizing these concepts for each approach in turn.

Decision Theory. For DT, the concepts of interest are those of value or utility; they are relevant only to judgments of preference. A value function represents a formalization of an individual's value or preference structure for decision problems under conditions of certainty; it serves to compare various levels of different attributes indirectly. In particular, a value function v associates a real number $v(x)$ to each point in an evaluation space, provided that, if the decision maker prefers x' to x'' (is indifferent between them), then he prefers $v(x')$ to $v(x'')$ (is indifferent between them). In essence, what the value function does is to rescale the objective values of an attribute X . A graphical representation of the v -function plots the x values, using the natural units of the attribute, on the ordinate and the $v(x)$ values, ranging from 0 to 1, on the abscissa.

In terms of procedures, after the range of the attribute is established, the fractile method described earlier is used to find the value X' such that the decision maker is indifferent between receiving a 50-50 lottery offering X_0 or X_1 , or receiving X' for sure. This mid-value splitting technique is iterated until sufficient points have been established to enable plotting of the function.

If the preferential independence assumption is met (see above), this same procedure can be extended to the n -attribute case. After value functions have been established for each attribute, the weights associated with each attribute are determined and the utility of each of the set of alternatives A_i , $i = 1, \dots, n$ is evaluated (see below for a discussion of the procedures involved in these steps).

Utility functions are not identical to value functions; a utility function is a value function, although the converse is not necessarily true. Utility functions are appropriate for the assessment of preferences under conditions of uncertainty. They are more complex to assess than value functions, since the decision maker's risk preferences must be taken into account.

The procedures for assigning utility functions to probabilistic alternatives are described below. We begin with the procedures for alternatives evaluated in terms of one attribute X . First, the ranges of the evaluation scale, from the minimum possible to the maximum possible values, are defined. The scale may be expressed either in the natural units of the attribute or in terms of a subjective scale. The decision analyst then identifies certain relevant qualitative characteristics of the utility function. He determines if the u -function is monotonic; i.e., whether for $x_1, 1, \dots, n$, it is always the case that x_j is preferred to (not preferred to) x_1 , where $j > 1$. He also determines the decision maker's attitude toward risk; i.e., he asks the decision maker if he prefers the certainty option offering x_1 or a 50-50 lottery offering x_{1+k} or x_{1-k} for a number of different levels of both x_1 and k . The point here is to identify the general form of the utility function; the form of the function will be concave if the decision maker is risk averse and convex if he is risk prone.

Next, the quantitative characteristics of the utility function are determined in an analogous fashion to that which was done with value functions. A few points on the utility function are fixed

using the fractile technique. That is, the procedure starts by identifying the certainty equivalent X' such that the decision maker is indifferent between it and a 50-50 lottery offering X_0 and X_1 (the best and worst possible outcomes). The procedure is iterated until a sufficient number of points have been fixed to plot a straight line describing the function, or to fit a parametric curve, if the decision maker is risk prone or risk averse. Finally, the expected utility of each probabilistic alternative is evaluated (see below for a discussion of the relevant procedures).

In the extension to the n -attribute case, the analyst is dealing with vectors rather than scalars. Utility must be assessed over a multiattribute consequence space. The procedures are essentially the same as above, however, if the utility independence assumption is satisfied (see above), for all pairs of attributes. If the utility independence assumption is not satisfied, the analyst will ordinarily attempt to redefine the attributes so they are utility independent or directly assess preferences within the multiattribute consequence space. After utility functions are defined, the scaling constants associated with each attribute are assessed and the expected utilities of the alternatives are calculated (see below).

Several comments concerning DT's procedures for assessing utilities are warranted here. These procedures are to a considerable degree a legacy of DT's theoretic origins in the von Neumann-Morgenstern and economic schools of thought. They are derived from earlier work involving objective probabilities and values. They

have been criticized by some (primarily within BDT) for being overly complex in comparison to more direct estimations of utility. Because of its theoretical origins and orientation, however, DT prefers to measure utilities in terms of preference and rarely makes use of the more direct types of techniques found within BDT. Although the procedures of DT are not, perhaps, as complex in practice as they may sound in description, they are in all likelihood more complex (and probably more time consuming) than the procedures used by other approaches.

Behavioral Decision Theory. For BDT, the concept of interest is that of the value curve. Although they share similar theoretical origins, DT and BDT differ considerably concerning the best procedures for assessing value curves. Although BDT occasionally uses similar techniques to those of DT in its basic research function, within its practical application function it assesses value curves in a much more direct manner.

The concept of value curve in BDT refers to an individual's value or preference structure for each dimension of value used in evaluating the relative desirability or utility of competing alternatives. The value curve, then, represents the functional relationship between an attribute value and judgments of desirability. In graphic terms, the x-axis represents the plausible range of a dimension (which is not necessarily numeric, e.g., it may range from "poor" to "excellent"), while the y-axis represents desirability (with the same metric for each dimension).

Procedurally, the first steps in BDT involve the identification of the appropriate decision maker(s), the alternatives to be

evaluated, and the relevant dimensions of value. Also, the relative weights associated with each dimension are identified (see below) prior to the specification of the value curves.

Value curves are ordinarily assessed by one of two methods. If the utility associated with an underlying dimension is conditionally monotonic, the BDT analyst often approximates the value curve by drawing a straight line. The justification for this simplified procedure for defining value curves is the robustness of linear approximation after aggregation across attributes. If the utility associated with an underlying dimension is not monotonic, or if the decision maker emphasizes the nonlinearity of his value function, the usual procedure is simply to ask the respondent to draw a graph describing his value curve for that dimension.

Psychological Decision Theory. PDT conceptualizes value and utility in a somewhat different fashion from DT or BDT. In prospect theory (Kahneman & Tversky, 1977), the carriers of value in problems of risky decision making are changes in wealth or welfare, rather than final states, and value is treated as a function of (a) the asset position that serves as a reference point and (b) the magnitude of the change (positive or negative) from that reference point. PDT, like IIT below, attends more closely than DT to the "psychological meaning" of value or utility. In general, the value function is concave for gains and convex for losses and is steeper for losses than for gains. In terms of procedures, the value curve for an individual is constructed on the basis of his choices among risky prospects of the form

"receive x with probability p , y with probability q , and (if $p + q < 1$), nothing with probability $1-p-q$."

In some of the PDT work involving regression techniques, the concept of functional relation also appears. This concept is essentially identical to the concept of function form in SJT and is operationalized and defined in basically the same manner. Since the concept is more central to the SJT approach, we turn to it.

Social Judgment Theory. In SJT, the concept of function form refers to the functional relationship between the values of a cue and either an individual's ratings of desirability or his inferences concerning the value of a criterion. The function form relates values of an attribute (in terms of the ranges appearing in the task) to values of the judgment (made on an interval scale); graphically, it is similar to the concepts of value or utility in DT or BDT. Unlike DT or BDT, however, the function forms associated with each dimension are all assessed simultaneously, along with the weights associated with each dimension.

In terms of procedures, the judgment problem and its relevant dimensions are first identified, and the ranges for each cue are established by the judge and/or the researcher. A number of hypothetical profiles or cases are then constructed (or, a number of real profiles selected). These profiles contain different combinations of the values of the cues and are presented sequentially to the judge, frequently via interactive computer techniques. The judge rates the desirability of each or estimates the value of the criterion, responding on an interval scale.

The relationship between judgments and cue values are then described by multiple regression statistical techniques. The basic model is a nonlinear additive one which includes squared cue values as well as the natural cue values. Inclusion of the squared terms permits the description of U and inverted-U shaped function forms or any cue-judgment function that can be approximated by some portion of a parabola, as well as linear function forms. In practice, SJT algebraically transforms the additive model into a weighted averaging model in which the relative weights for each attribute sum to one and the function forms share a common range. Function forms are ordinarily displayed graphically (along with relative weights) via interactive computer graphics techniques. Alternatives may then be assessed via application of the weighted averaging model (see below). Judges are sometimes asked simply to draw the desired function form for a cue, especially after a preliminary analysis based upon their wholistic ratings of multiple cases.

Information Integration Theory. IIT can be distinguished from DT, BDT, and SJT in that it attends to the subjective scaling of attributes. That is, rather than relating judgments to the natural metric of the dimensions, IIT relates judgments to the subjective metrics of the dimensions; objective values are translated into scale values.

The primary concerns of IIT are with the algebraic models that best describe human information integration and with substantive theory. Scales are of secondary interest in so far as the functional scale values of stimuli are dependent upon the model best describing the judge's information integration processes. The central issue in

IIT theory, then, is the test of the goodness of fit of a model to the data. Such a test provides joint validation of the model and the response scale. In other words, the scale value of stimuli can only be discussed in terms of an integration model with good fit to the data.

Procedurally, the judgment analyst's first constructs stimulus combinations, using factorial design. The judge then rates each combination, on a scale assumed to possess interval properties; this assumption is tested within the course of IIT procedures. The judge may express his rating in terms of numbers, lengths of lines, graphs, and so forth.

A test of the goodness of fit of the model to the data is then performed. For example, if the hypothesized model were an averaging model, the test of goodness of fit would require no significant interactions in an analysis of variance and parallelism when levels of each stimulus were plotted as a function of the levels of each other stimulus. If the goodness of fit test is passed, the response scale is validated as an "equal-interval" scale. Satisfactory fit testifies to the adequacy of the response measure, and such validation of the model thus provides a functional scaling of the response. Observed row means then provide estimates on an interval scale of the scale values of the row stimuli.

The same general procedure is followed for the tests of adding, subtracting, or multiplying models. If the model satisfactorily passes the appropriate goodness of fit test, then row means can be interpreted as scale values of the stimuli. Many IIT studies,

however, provide little information concerning scale values because of the small number of factorial levels involved in the design. For example, a 2 x 2 ANOVA design provides essentially no information on the stimulus scales, since the 0 and 1 points on a scale are arbitrary.

Attribution Theory. Only a weak form of a concept of functional relationship is found in the AT approach, that of directionality. Directionality indicates the form of the relationship between levels of attributes and levels of judgment. Within DT, these descriptions are usually nomothetic and are frequently based upon the results from analyses of data from ANOVA designs. For example, it is common to encounter statements of the form: "Low levels of attribute X are associated with judgments that the level of criterion Y is low, while high levels of attribute X are associated with judgments that the level of criterion Y is high."

Procedurally, a number of stimulus combinations are first generated, usually in terms of factorial design and frequently involving only two levels of each attribute. The judges then rate some or all of the attribute combinations; usually, each judge responds to only one or a small subset of all possible combinations. Data is ordinarily aggregated across judges; for those attributes that statistical analyses (usually, ANOVA) indicate have a significant effect on judgments, the direction of the effect is reported.

Overview. All six approaches contain some concept related to the functional relations between attributes and judgments, although there exists considerable variation in both the nature of these concepts and, particularly, in the procedures by which they

operationalize those concepts. The end output of the procedures of DT, BDT, PDT, and SJT, however, seem much the same. Each of those approaches generates a description of the relationships between objective attribute levels and individuals' judgments of desirability (or of the value of a criterion) and this description is frequently represented in graphic form. AT also commonly describes such relationships between attribute levels and judgments, although in less specific form and usually in reference to groups of judges rather than individuals. IIT takes the most divergent approach to this issue, rescaling objective attribute levels into psychological scale values. Formally, however, the end result of the functional measurement approach does not seem greatly different from those of the other approaches (see, Shanteau & Phelps, 1977).

The major procedural difference among the approaches is probably the manner in which they elicit responses from judges in order to identify utility or value functions, function forms, or scale weights. The procedures can be divided into three general categories. DT and PDT (and, occasionally, BDT) rely upon procedures involving the indication of preference between lotteries, bets, or other risky choices. BDT usually relies upon more direct measures of identification or construction, approximating value curves with straight lines or simply having the judge draw the value curve; at times, SJT also makes use of this latter procedure. SJT, IIT, and AT customarily require the judge to make wholistic, intuitive ratings.

Two issues concerning the relative merit of the various procedures for identifying the functional relationships between attribute

values and individual's judgments are deserving of note. The first of these is the issue of ease of use and practicality. Under this heading would also fall such considerations as user satisfaction and time required for the conduct of the procedures. Although empirical tests are required to establish the facts of the matter, the procedures of the DT approach appear to rate lowest in this respect because of the number of questions to which the decision maker must respond and the amount of time required for the specification of value or utility curves. On the face of it, the procedures used by BDT in practical applications appear to be the simplest and least time consuming.

The second issue is the type of error to which the procedures are most prone or resistant. In particular the measurement error associated with any procedure can be divided into two components: errors of estimation and errors of elicitation. Estimation error is error associated with imprecision of procedures from a technical standpoint; these can be described as "errors in principle." Elicitation error is error associated with more psychological factors, such as fatigue, boredom, or inattention; these can be described as "errors in practice." This second issue is not unrelated to the first. DT, for example, with its rigorous and exhaustive procedures for testing independence assumptions and constructing value and utility functions would seem to be on one end of the continuum, insofar as it employs procedures that are designed to minimize errors of measurement. BDT, on the other hand, with its simple and direct procedures appears to be on the other end of the continuum,

with procedures designed to minimize errors of elicitation. Which type of error is more important in the analysis of judgment or decision processes is, of course, a topic for empirical research.

Weight. All six approaches contain some concept related to weight, or the importance of a piece of information to individuals' judgments of preference or inferences about states of the world. We discuss in turn the procedures used by each approach in operationalizing the concept of weight.

Decision Theory. DT has two concepts related to weight. The first is that of scaling constants which are used to weight the various unidimensional value or utility functions that enter a multidimensional utility function. The second is that of subjective probability which is used to weight the utility of an outcome when making a decision under conditions of uncertainty; that is, subjective probabilities are used to weight utilities in generating estimates of the subjective expected utilities of probabilistic alternatives.

We turn first to the concept of scaling constants and to the procedures for assessing such constants for multiattribute value functions under conditions of certainty. We start with the simplest case, that involving only two attributes. First, the unidimensional value functions for two attributes X and Y are obtained using the procedures described above. The analyst then asks the respondent to make a series of choices between alternatives that will permit him to calculate the two scaling constants k_x and k_y . In order to do so, two (X,Y) pairs to which the decision maker is indifferent must be identified.

The analyst begins by asking the decision maker which of the pairs (X_0, Y_1) and (X_1, Y_0) is preferred (where X_0 and Y_0 are the worst levels of those attributes and X_1 and Y_1 are the best). The decision maker can indicate that he is either indifferent, prefers the option with the highest X value (X_1, Y_0) , or prefers the option with the highest Y value (X_0, Y_1) . If he is indifferent, then, $k_x = k_y = .5$; if X is preferred, then, $k_x > k_y$; and, if Y is preferred, then, $k_y > k_x$. Supposing that X is preferred to Y, the analyst then asks the decision maker a series of questions designed to identify a value, X_1 , such that he is indifferent between (X_1, Y_0) and (X_0, Y_1) . Knowing the value X_1 , and given that $\sum_{i=1}^n k_i = 1$, k_x and k_y can then be derived using simple algebra. This procedure can easily be generalized to the case of more than two attributes.

We now turn to the procedures for assessing the scaling constants for multiattribute utility functions under conditions of uncertainty. Again, we deal with the case of two attributes. The form of a multiattribute utility function can be specified in terms of a number of conditional utility functions over the attributes and scaling constants. For example, a multilinear utility function (the most general form) with two attributes can be specified as

$$u(y, z) = k_y u_y(y) + k_z u_z(z) + k_{yz} u_y(y) u_z(z),$$

where there is one conditional utility function for each Y and Z, and three scaling constants: k_y , k_z , k_{yz} . The two utility functions u_y and u_z are scaled from 0 to 1.

The basic idea underlying the procedures is to obtain a set of three independent equations with three unknowns, which are then solved to find the k s. The procedures for generating these equations may involve either certainty comparisons or probabilistic techniques. A typical set of procedures might involve the following steps:

If we assume that preferences are a monotonically increasing function of Y , Z and (Y,Z) , then the first equation can be easily generated; for (Y_1, Z_1) :

$$1 = k_y + k_z + k_{yz}$$

Then, using the techniques described above for the estimation of scaling constants under certainty, if the decision maker's preferences indicate that $k_y > k_z$, a value of $Y(Y_1)$ can be formed such that the decision maker is indifferent between (Y_1, Z_0) and (Y_0, Z_1) . The second equation can thus be written as

$$k_z = k_y u_y(Y_1).$$

Next, using probabilistic scaling techniques, a third equation can be generated via a series of questions designed to find the probability, (p_y) such that the decision maker is indifferent between the consequence (Y_1, Z_0) for sure and a lottery offering consequence (Y_1, Z_1) with probability, p , and consequence (Y_0, Z_0) with probability $1-p$. Equation 3, then, can be written as

$$k_y = p_y$$

These three equations can then be solved for the values of the scaling constants.

These procedures can be generalized to the case of more than two attributes. In the n -attribute case, however, the rank order of the values of the k s are determined before their magnitudes are assessed. The customary procedure for this ranking involves setting the values of each attribute at their lowest levels. The analyst then asks the decision maker to indicate which of the attributes he would move to its highest level, if he could move only one. Then, he is asked which of the remaining attributes he would raise to its highest level, and so on, until the rank order of the importance of each attribute has been established.

The second concept in DT that sometimes serves as a weighting factor is that of subjective probability, the assessment of which we have already described. In evaluating multiattribute utility functions under uncertainty, the utilities of probabilistic alternatives must be transformed into subjective expected utilities in order to make the final evaluation of an alternative. This transformation is accomplished by weighting utility by the subjective probability of the event. If the estimate of probability is a point estimate, the procedure for calculating subjective expected utility is given by the formula, $SEU_i = p_i u_i$, where SEU is the subjective expected utility of an alternative, p_i is its probability, and u_i is its utility. If the subjective probability of an alternative is expressed in terms of a probability distribution, the procedure for determining its subjective expected utility is given by the formula $SEU_i = \int p_i(u_i)$, where fp_i is the probability distribution for the i th alternative.

Behavioral Decision Theory. Within BDT, there also exist two concepts related to weight. The first is that of weight, which refers to the relative importance the judge assigns to relevant dimensions in evaluating preference for multiattribute alternatives. Procedurally, after the entities to be evaluated and the relevant dimensions of value have been identified, the judge ranks the dimensions in order of importance. He then rates the dimensions in order of importance, preserving ratios, with the least important dimension assigned a rating of 10. The ratings are then summed and each is divided by the sum in order to establish a measure of the relative importance of each dimension; relative weights, therefore, sum to one. Competing entities can be evaluated according to the formula $U_i = \sum_{j=1}^n w_j u_{ij}$; i.e., for each alternative, its overall utility is the weighted sum of the utility of each attribute level for that alternative.

The second concept is that of subjective probability. In basic research in BDT, it is sometimes used in the same manner as in DT to calculate the subjective expected utility of probabilistic options or alternatives.

Psychological Decision Theory. In PDT, the concept of decision weight replaces that of subjective probability in the analysis of decision making under risk. According to prospect theory, risky options are evaluated by the formula, $V_i = w(p_i)v_i$, where the decision weight, $w(p_i)$, replaces the probability p_i used in utility theory and the value function, v_i , (the procedures for assessing v_i are discussed above) replaces the utility function of DT or BDT. Decision weights are generally

lower than the corresponding probabilities, except in the range of low probabilities. In terms of procedures, decision weights are assessed by the same type of procedures used by DT or BDT to assess subjective probabilities. That is, they are inferred from preferences between risky options, or prospects. Decision weights do not obey the probability axioms, however, nor do they provide an adequate measure of degree of belief.

The concept of weight also appears in some of the work in PDT employing regression techniques. In this component of PDT, weight refers to the relative importance of attributes in predicting judgments of preference or inference. It is operationalized in terms of the relative magnitude of beta weights.

Social Judgment Theory. A similar concept of weight appears in SJT. In SJT, the concept of weight refers to the relative importance the judge places on cues in making judgments of preference or inference. Operationally, the relative magnitude of the beta weights associated with each cue are scaled so that they sum to one. The procedures involved in assessing weights are identical to those described for assessing function forms, since the two are assessed simultaneously. In particular, the procedures are based on intuitive, wholistic ratings of a number of hypothetical (or real) profiles. Relative weights are ordinarily displayed graphically in terms of bar graphs, the relative lengths of which indicate the importance of each dimension, as well as by a numeric value between 0 and 1. Relative weights are combined with function forms to evaluate individual profiles in

terms of the formula, $Y_i = \sum_{j=1}^n w_j v_j(x_{ij})$, where the value of the judgment, Y_i , is the sum of the products of the relative importance of each cue (w_j) times the value of the cue level specified by the function form ($v_j(x_{ij})$).

Information Integration Theory. In IIT, the concept of weight refers to the salience or importance of the information contained by a stimulus item. It may be regarded as the amount of information in the stimulus. The concept of weight is operationalized in IIT differently from the other approaches. All the other approaches derive weights from the data; they normally make no a priori assumptions about the values of weights, other than restrictions such as that the weights must sum to 1. IIT, however, specifies weights a priori in terms of the particular model of the information integration process to be tested. Weights are ordinarily assumed to be equal across the stimuli, unless there are theoretical reasons to set them otherwise a priori. For example, stimuli might be assigned greater or lesser weight according to their serial position in a sequence. Relative weights can, of course, be manipulated by varying the number of equivalent stimuli in a subset, or by manipulating the reliability of a piece of information. Specific procedures for estimating differential weights are fairly complex and nongeneral (see Anderson, 1972). Because they are used infrequently, they will not be discussed here.

In IIT studies of risky decision making, the probability of events is assumed to modify value in determining the subjective utility of bets. Such probabilities, therefore, can also be thought

of as weighting factors. The treatment of probabilities by IIT is similar to the decision weight concept in PDT's prospect theory. The "weight" associated with a probability is not a direct function of its magnitude. Rather, the scale value of the probability (as determined by the procedures described above) enters the assessment of probabilistic bets. That is, the utility of simple bets is a function of the scale value of the given probability times the scale value of the outcome.

Attribution Theory. AT contains no well-articulated concept of weight, particularly with respect to individual judges. A related idea appears, however, in nomothetic statements concerning the judgment processes of groups of judges. This idea is usually presented in a dichotomous yes/no form and is based upon standard tests of statistical significance. In particular, statements of the type "attribute X significantly affects (does not significantly affect) judgments about criterion Y" are frequently found within AT.

Overview. All six approaches contain concepts referring to the importance of particular attributes in judgments of preference or inference. A major topic over the past few years has been which set of procedures leads to the most accurate descriptions of such weights. Recently, however, a number of theoretical and empirical studies (Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975; Schmidt, 1971; Wainer, 1976) have demonstrated that in many instances all possible sets of weights produce essentially the same results. The conclusion suggested by such research is that the establishment of differential weights is unnecessary; equal weights will serve just as

well. Recent theoretical research by McClelland (1978) indicates, however, that the determination of optimal differential weights in multiattribute decisions is more commonly important than the results of the above studies had indicated. In any case, there appears to be no empirical evidence to indicate that any one of the six approaches' procedures are significantly better than any other in eliciting optimal weights.

The major issue concerning the relative merit of various procedures for eliciting weights, therefore, may not be their relative accuracy, but rather more practical and psychological issues such as ease of use, user satisfaction, and time required. Again, the DT approach appears to be the most complex and time consuming when a large number of scaling constants must be assessed. They appear to be relatively simple and quick, however, when only a small number of attributes are involved in the multiattribute utility function. There does not appear to be great differences among the other approaches with respect to ease of use, satisfaction, time required and so forth. Empirical research is necessary, of course, to provide reliable evidence concerning these issues.

One procedural difference among the approaches which has yet to receive empirical attention, but which might prove of interest is the order in which weights and utility functions or function forms are assessed. Ordinarily, utility curves are assessed first, then weights in DT; weights are assessed first, then value curves in BDT; functional relations and weights are assessed simultaneously in PDT, SJT, and IIT. There is some reason to believe that the order in

which these parameters are assessed might affect the obtained values (Adelman, Note 1; Lonsdale, Note 2).

Finally, it is interesting to note that IIT and the other approaches differ in the methods and procedures by which they indicate the relative importance of an attribute or stimuli in the individual's judgment or decision processes. DT, BDT, PDT, and SJT ordinarily scale functional attribute-judgment relations over a common range and let the value of the weights associated with those attributes vary. IIT, however, ordinarily holds weights across stimuli constant and varies the subjective value of the stimuli. Although there are considerable and significant theoretical differences associated with this difference in approach, procedurally, the two techniques seem to accomplish more or less the same end.

Organizing principles. The next major issue is the manner in which individuals organize or integrate information into a judgment or decision. The various approaches have used several classes of models to describe individuals' organizing principles.

Bayes theorem. The first class is typified by Bayes theorem, which specifies the optimal means of revising objective probabilities. Within DT, Bayes theorem is prescribed as the appropriate means for revising subjective probabilities, and it is used on those occasions in practical decision problems when subjective estimates of probability must be revised or integrated.

BDT has also made extensive use of Bayes theorem in its basic research. Much of this research has involved tests of the adequacy of the theorem as a model of individuals' revisions of subjective

probability. Similarly, there is a substantial body of research within PDT which involves tests of the adequacy of Bayes theorem (and similar optimal models for probability revision or estimation) as a descriptive model of the organizing principles used by judges in integrating probabilistic information. In terms of procedures, such tests have customarily involved the elicitation of estimates of probability on the basis of a given set of information. Such estimates are then compared to those generated by Bayes theorem; the degree of congruence between the estimates elicited from individuals and those generated by Bayes theorem indicate the adequacy of the theorem as a descriptive model of individuals' organizing principles.

Additive models. The next class of models are additive models for combining and integrating the relative weights of attributes and the functional value of attribute levels into an overall judgment. Simple additive models with no interactive terms are described by the following formula:

$$Y' = \sum k_i u_i(x_i)$$

where Y' is the predicted or prescribed response, k_i are the weighting factors associated with each attribute, and $u_i(x_i)$ are the values or utilities associated with the particular levels of the i attributes.

DT uses this model when the sum of the scaling constants equals one ($\sum k_i = 1$). Procedurally, the scaling constants are first established by setting up a series of equations, as described above, without the restrictive assumption that $\sum k_i = 1$. The equations are

solved for the n k -values; if $\sum k_1 = 1$, the analyst then asks a series of questions designed to find whether there are any interactions between levels of specific attributes that result in non-additivity for the decision maker. If no such interactions are indicated by the decision maker's responses, the additive model is used. The various alternatives can then be evaluated via application of the above formula.

Although the above model is described by DT as an additive one, it can also be described accurately as a weighted averaging model because of the restriction that the scaling constants sum to one. The restriction that the scaling constants sum to one ensures that the relative weights associated with attributes (can) change as attributes are added to or deleted from the utility function. Weighted averaging models are thus a special case of additive models.

BDT, in its applied version, exclusively relies upon weighted averaging models in describing a decision maker's value function. That is, the model includes no multiplicative terms and the relative weights sum to one. BDT justifies reliance on weighted averaging models on the grounds that even when the assumptions for such models are not satisfied, they ordinarily provide good approximations to the solutions provided by more complex models. Competing alternatives are evaluated in terms of a weighted averaging model in order to aid the decision maker in making a choice for action.

SJT also relies on weighted averaging models to describe individuals' judgment and decision processes, although occasionally making use of multiplicative, configural, or other noncompensatory

models. SJT is perhaps the best example of an approach in which concepts concerning the nature of the objective characteristics of the task environment affect the concepts and procedures associated with a judge's organizing principles. SJT holds that environmental task systems can ordinarily best be described in terms of a weighted averaging model that involve the association of relative weights and function forms with each attribute; parallel concepts are used to describe the organizing principles of individual judges. In SJT, weighted averaging models may be used either to evaluate competing alternatives or to describe judge's judgment or decision processes; that is, they may serve either prescriptive or normative functions.

The distinction between additive and averaging models is an important one in IIT. Indeed, IIT has conducted considerable research concerning the adequacy of additive versus averaging models as a description of the information integration process. The averaging model in IIT can be described in the following manner:

$$R = C + \frac{\sum_{i=0}^n w_i s_i}{\sum_{i=0}^n w_i}$$

where C allows for an arbitrary zero in the response scale, w_i are the weights associated with each stimulus, and s_i is the scale value for each level of the stimuli with s_0 as an internal organismic variable. This formulation has the effect of scaling weights so that that they sum to one. An additive model in IIT can be represented by the above formula with the omission of the denominator term. It might be described more precisely as a summative model. Research within IIT indicates that judges most often act as if they were

following the above averaging model. For example, in studies of impression formation, judges have given higher responses to high-high adjective combinations, as predicted by an averaging model, than to high-high-medium-medium combinations, as predicted by a simple additive model.

Multiplicative models. The next class of model commonly found within the six approaches is that of multiplicative models. The general form of this model is given by the equation

$$Y = \sum w_i f_i(x_i) w_j f_j(x_j)$$

where Y is the predicted or prescribed response, w_i and w_j are the weights associated with the i^{th} and j^{th} attributes, and $f_i(x_i)$ and $f_j(x_j)$ are the values or utilities associated with the particular levels of the i^{th} and j^{th} attributes. This model allows for configularity and weak interactions among attributes.

DT uses this model in multiattribute utility functions when it is found that the decision maker is not indifferent between a lottery offering (x, y) or (x', y') with equal probability and a lottery offering (x, y') or (x', y) with equal probability, where $u(x) \neq u(x')$ and $u(y) \neq u(y')$, and all other levels of attributes are fixed.

Multiplying models are also frequently found in IIT. For example, Shanteau (1975) applied an IIT approach in an analysis of risky decision making. For one part gambles of the type "P to get \$," where P is the probability of getting outcome \$, the subjective value of \$ was represented by its subjective scale value, s , and the subjective likelihood, P, of receiving the outcome was represented by its weight value, w , reducing the information integration

model to a multiplicative one, $R = ws$. Shanteau reported that such a multiplicative model described well individuals' preferences for one part gambles.

Multiplicative models are also used, of course, in the basic research of BDT and PDT on subjective expected utility and choice under uncertainty. In BDT, subjective expected utility is calculated by multiplying subjective probability times utility. In PDT, the attractiveness of risky choices is evaluated by multiplying decision weights times the value of the gains or losses entailed by an outcome.

Multilinear models. Multilinear models involve joint adding and multiplying functions. Formally, they are, of course, more general formulations of both adding and multiplying models. Complex multilinear models employing both multiplicative and additive components are found in both DT and IIT approaches; they are not commonly found within any of the other approaches.

Schemata. The final class of models is that of schemata, which appear in AT. According to Kelley (1973), schemata are "assumed patterns of data in a complete analysis of variance framework" (115). These schemata can be interpreted as a repertoire of abstract ideas about the operation and interaction of causal factors. The individual uses these schemata to fit bits and pieces of relevant information into an interpretable framework and, thus, to reach (usually reasonably good) causal attributions.

Two major principles guide the attributional analysis, as well as the construction of causal schemata. These are the

covariation and discounting principles. The covariation principle states that an effect is attributed to the one of its plausible causes with which it covaries over time. The discounting principle states that the role of a given cause in producing an effect is discounted if other plausible causes are also present.

In short, causal schemata are conceptions of the manner in which two or more causal factors interact in relation to a particular kind of effect. They can be described as similar to an analysis of variance layout; the individual may "fill in the cells" either by observation or by "thought experiments." The judge relies upon these schemata in organizing information into inferences about the causes of events.

In terms of procedures, the adequacy of schemata as models of individuals organizing principles is tested via traditional hypothetico-deductive techniques. Predictions of judges' inferences are derived from these models and the responses from judges are tested against those predictions. Congruence between predictions and responses are interpreted as confirmation in support of the models as a description of the organizing principles used by judges in making inferences.

Other models also occasionally find their way into the procedures of the six approaches. Among others, these include models based upon conjoint measurement techniques; disjunctive, conjunctive and other noncompensatory models; and, models formulated only in verbal terms. An exhaustive cataloging of all the various models that have appeared in the procedures of the six approaches, however, would be quite lengthy and is beyond our purpose here.

Overview. The six approaches use a variety of models in describing or prescribing the organizing principles involved in individuals' responses to judgment or decision problems. The most commonly employed models can be placed into three categories. These models can be distinguished further according to their intended aims or purposes.

The first of these three categories is best represented by Bayes theorem and includes other models that specify the "objectively" optimal means for integrating or organizing information. Within DT, Bayes theorem is used as a prescriptive aid for decision makers. That is, it is used to prescribe the manner in which the decision maker should revise probabilities on the basis of new information or integrate probabilistic information from multiple sources. Bayes theorem and other prescriptive models have been used as hypothetical models of the organizing processes individuals use in integrating probabilistic information in research within BDT and PDT; the results of such tests of Bayes theorem as a descriptive model have been mixed.

Additive, weighted averaging, and multiplicative models can be categorized together as special cases of multilinear models. DT makes use of all the models in this second category; these models describe the decision makers preference or utility structure, but their primary purpose is prescriptive. That is, they are used to evaluate competing alternatives in making a choice for action. BDT relies upon weighted averaging models for the same purpose. That is, weighted averaging models are used to help the decision maker

formalize his value structure and to evaluate competing alternatives. DT, BDT, and PDT all use multiplicative models in descriptions of the integrating processes involved in the evaluation of risky alternatives.

Although other models occasionally appear, SJT customarily relies on weighted averaging models for both descriptive and prescriptive purposes. Historically, these models were first developed to describe the organizing principles used to integrate information from multiple sources into inferences or judgments. More recently, such models have been used in a somewhat more prescriptive function; that is, to aid decision makers in making choices for action.

IIT makes use of all models in the second category, but their intended function is always descriptive. IIT seeks to find models that adequately fit the data describing judges' use of information.

The final category of models is typified by the schemata of AT. These models are always descriptive in function. They are generally not as precisely or quantitatively formulated as models in the above two categories.

[MORE]

Psychological factors. All six approaches contain concepts referring to psychological factors that bias or otherwise influence judgment or decision making processes. PDT is perhaps most prominent in this respect, including, for example, concepts such as representativeness, availability, anchoring, causality, and diagnosticity (see Theory section for a more detailed discussion of these concepts) which refer to factors that distort the judgment or

decision making process and lead to less than optimal performance. AT also contains a number of concepts referring to such biases, for example, the so-called attribution error, which refers to the tendency of individuals to overattribute effects to personological, versus environmental, causes. These concepts do not ordinarily affect procedures directly, although PDT has specified corrective procedures for such biases in applied decision making contexts (Kahneman & Tversky, 1978).

In exception to the general state of affairs, SJT routinely operationalizes and measures one of its major concepts relating to psychological factors, that of cognitive control. Procedurally, it can be viewed as a test of a particular model with the data, similar to IIT's goodness-of-fit tests; it can be operationalized as the correlation between observed judgments and those predicted by a given model. More specifically, it can be interpreted as variance in judgments "around" a given model. If the given model is the "best" model of an individual's judgment processes, cognitive control can be interpreted as a measure of the ability to exercise or apply a given model consistently, without error.

Less than perfect correlation between predicted and observed judgments can be indicative of either less than perfect cognitive control or an ill-fitting model. The correlation between repeated trials is indicative of the source of error. To the extent the correlation between repeated trials, or consistency, is high, a less than optimal model is indicated. To the extent consistency is low, less than perfect cognitive control is indicated as the source of error.

[MORE]

Overview. The discussion of the subjective use of objective data, cue utilization, and central processes loci has focused on four topics: (a) utility/value/function forms, (b) weights, (c) organizing principles, and (d) psychological factors. All of the approaches contain important concepts related to each of these topics, although they show considerable variety in the procedures they use in operationalizing those concepts. It does not seem a drastic over-generalization to conclude, however, that the procedures of all six approaches are directed toward describing or prescribing judgment or decision processes that include: (a) the evaluation of levels of individual attributes and/or of interactions between levels of attributes, (b) the association of levels of relative importance (weights) to each attribute or interaction, (c) the aggregation or integration of such evaluations and weights into an overall judgment, and (d) psychological factors that may affect the above processes.

[MORE]

G. Task-Judge Interaction (Table IV-G)

We turn finally to those concepts related to the interaction of the judge or decision maker with the task environment. For some of the approaches, these concepts refer only to the interaction between the individual and the task "surface"; that is, they refer to the relation between the judge and those characteristics of the task that directly impinge on him. Other approaches include, in addition, concepts referring to the relation between the judge and

Table IV-G

| DT | BDT | PDT | SJT | IIT | AT |
|--------------------|-----|-----|---------------------|-----|----|
| <u>Optimality</u> | | | | | |
| consistency | X | X | (cognitive control) | | |
| | X | X | accuracy | | X |
| | | | knowledge | | |
| <u>Feedback</u> | | | | | |
| | X | | outcome/cognitive | | X |
| <u>Feedforward</u> | | | | | |
| | | | X | | |

the task "depth"; that is, they refer to the relation between the judge and those characteristics of the environmental system which he cannot directly observe, but must infer on the basis of those task elements with which he can directly interact. Stated more simply, the concepts of the first type are not concerned with "correct answers," whereas those of the second type are.

Optimality. The most general concepts concerning task-judge interactions involve the idea of optimality. The optimality (or suboptimality) of individuals can be described either in reference to their relation with surface elements of the task, their relation with depth elements, or both.

The concept of consistency refers to the relation between the judge and the surface elements of the task. It is applicable in those situations in which individuals must make judgments of preference, i.e., in those situations in which there are no right or wrong answers. Optimal behavior in such circumstances can therefore be measured only in terms of internal consistency across multiple instances in integrating information into a judgment or choice.

DT handles the optimality problem by formalizing decision makers' preference and utility structures and using these formal models to evaluate multiple alternatives. The evaluation of alternatives is thus optimal in the sense that decision makers' preference and utility structures are uniformly and consistently applied to each alternative.

BDT, in its application function, deals with the problem in the same manner as DT. BDT also often uses the concept of consistency

in basic research on preferences among risky options. Here, consistency refers to logical consistency, usually with respect to the transitivity or intransitivity of choices.

The same concept of consistency appears in PDT. Indeed, much of the research within PDT has been directed toward discovering and describing the psychological factors that lead to logical inconsistencies in choices among risky options. PDT has argued, however, that many of the apparent inconsistencies in preferences between risky options do not indicate suboptimality on the part of the judge, but rather inadequacies in utility theory as a descriptive model of decision making under risk. Prospect theory has been developed in an attempt to better account for commonly occurring apparent inconsistencies in preferences among risky options.

Optimality with respect to surface task characteristics is addressed by the concept of cognitive control in SJT. This concept was discussed under "psychological factors" in the previous section, but it is also relevant here since it refers to the consistency with which individuals integrate information into judgments over multiple instances of a particular judgment problem. Moreover, a significant fraction of the basic research within SJT has been concerned with discovering types of task characteristics (e.g., task unpredictability) that lead to less than perfect cognitive control. In practical applications, SJT sometimes deals with the optimality problem by substituting models of the judge for the judge himself in the evaluation of alternatives, thus ensuring that such evaluation is carried out in a consistent manner.

Finally, neither IIT or AT contain well defined or frequently operationalized concepts with respect to the optimality of judges' interactions with the surface elements of task systems.

The optimality of judges' interactions with respect to the depth characteristics of the task is described in several of the approaches by the concept of accuracy, where accuracy refers to the congruence between an individuals' inferences and the correct answer. In BDT and PDT's studies of probability estimation, the concept of accuracy is often operationalized as the degree of correspondence between the individual's probability estimate and the estimate provided by Bayes theorem or some statistical model for integrating probabilistic information.

In SJT, accuracy in learning studies is operationalized as the correlation between the value of the correct answer (given by the task system) and the judge's inference about the value of the criterion. A more central concept in SJT concerning optimality is that of knowledge. Knowledge refers to the correspondence between the predictable components of both the environmental system and the judge's cognitive system. This concept often replaces that of accuracy because it is a better measure of optimality when the environmental system is not perfectly predictable. That is, when the environmental system is not perfectly predictable, the accuracy index cannot obtain unity, even when performance is optimal. The knowledge index eliminates the effects of imperfect predictability on the measure of optimality; it can be operationally defined as the correlation between the values predicted by the model best

describing the environmental system and the values predicted by the model best describing the cognitive system of the judge.

AT is concerned on a theoretical level with the accuracy of the individual's inference with respect to the environmental system, but, in terms of procedures, it ordinarily does not construct tasks in which the correct response is known. The concept of accuracy is, therefore, ordinarily not operationalized in AT. No concept of accuracy is commonly operationalized in DT or IIT.

Feedback. In studies of learning, the concept of feedback is an important one. It refers to information provided by the task about the quality of the judge's inferences. Of the six approaches, however, only BDT, SJT, and AT devote much attention to learning, and the concept of feedback is operationalized only within BDT and SJT.

BDT has conducted probability estimation studies in which the effects of feedback on judge's accuracy have been investigated. Procedurally, after the judge makes a probability estimate (either specifying it directly or indicating it indirectly by choosing from among a set of possible events or hypotheses that which he considers most likely), he is told the correct answer.

SJT also makes use of this type of procedure, describing it as outcome feedback. That is, after every inference about the true value of a criterion, the judge is given the correct answer. SJT distinguishes between outcome feedback and cognitive feedback. Cognitive feedback differs from outcome feedback in that it provides more than yes/no or correct/incorrect information to the judge. In particular, it consists of information describing the relationships

between attributes and the criterion (or between attributes and the judge's inferences). The most common procedure for giving cognitive feedback involves the judge first making wholistic inferences on the basis of a number of cases. Afterwards, he is shown graphic models describing the relationships between attributes and criterion values and relationships between attributes and his inferences. For the environmental system, the model shows the weights associated with each attribute, the functional relations between attributes and the criterion, and the predictability of the system. For the judge, the model shows the weights he associates with each attribute, the functional relations between the attributes and his judgments, and his degree of cognitive control. SJT has conducted numerous studies comparing the effects of outcome versus cognitive feedback on learning; the conclusion has been that cognitive feedback is generally superior in facilitating learning.

Feedforward. The concept of feedforward also appears in SJT. In essence, feedforward consists of information describing the relationships between task attributes and the task criterion which is given the judge before he interacts with the task. Procedurally, feedforward is usually displayed in the same manner as the cognitive feedback described immediately above.

Overview. Most of the six approaches contain concepts related to the degree of optimality of the individual's interaction with the task system. These concepts may refer to the individual's optimality with respect to those task elements which are directly observable (in DT, BDT, PDT, SJT) or to his optimality with respect to those task elements which he must infer (in BDT, PDT, and SJT).

The concepts of feedback from the task to the judge appears in BDT and SJT; the concept of feedforward appears in SJT.

H. Conclusion

We have attempted in this section to describe as completely and fairly as possible the major concepts and procedures found within the six approaches. It is with a great deal of trepidation that we submit this effort for the reader's evaluation. Its shortcomings are nearly as apparent to the authors as they surely are to the reader. Nonetheless, it seems to us important that such a description be developed. We hope that this has been a step in the right direction and that it will be improved in the future to the point that it becomes genuinely useful.

Epilogue

The Colorado Report served as a focus for a conference on judgment and decision making held at the University of Colorado on April 28-30, 1978. Thirty researchers and practitioners prominent in the field attended the conference. In this epilogue we briefly describe what happened at the conference, what integration we think was achieved, and what we propose to do next.

We asked a proponent of each of the approaches considered in the report to prepare a formal critique of our work. Because the proponent for SJT (Ken Hammond) participated in the preparation of the report, there are only five critiques. These critiques were distributed to all participants at the conference and, together with additional comments from each proponent, they formed the basis for discussion on the first day of the conference. The second and third days were devoted to discussions of specific issues raised in the report. These included "Equal weights: Do methods and procedures make a difference?"; and "What is a good decision, judgment, inference? How would you know one if you saw one?"

Although we have not attempted the impossible task of summarizing the discussions, we have included the five critiques of the Report in the Appendix so that the reader may see the major themes that concerned the proponents of the various approaches. Whereas the critiques led us to correct minor errors in the report (e.g., omitted or incorrect references), we have made no substantive changes to the text in response to them. For although we do agree with many of the

criticisms and suggestions for improvement contained in the critiques, and we are gratified that our critics took the time to review our attempt at an integrative framework and to offer their advice, it was our view that it was more useful to place the original manuscript and the critiques of it before our audience now, rather than to delay its availability until we could incorporate their substance into the report. Many of their suggestions will be incorporated in a revision of the report which will be published as a book during the coming year.

There was a certain amount of uneasiness at the conference about what was meant by the concept of integration. If our goal was the immediate construction of a single theory in which all six approaches would be included, then many participants were not sure that they wanted to help. Although the creation of the most parsimonious theory for judgment and decision is an important goal, the time for such a fully integrated theory is apparently yet to come. However, we do believe that now is the time to begin the communication and interaction between approaches that will make an attempt at a theory feasible in the future.

Our immediate concern is what should now be done to further the integrative effort. There is much speculation in the report, in the critiques, and in the conference discussions about how different approaches might be applied to the same problem. We therefore intend to ask representatives of each approach to apply their theories, methods, and procedures to one or two common problems. The set of

proposed solutions for a common problem should further illuminate the similarities and differences between the approaches which were the subject of this report.

The methods by which the representatives of each approach deal with the common problems posed by us will form the basis of our next conference and report. We hope that in the epilogue we write for that report we will be able to claim further progress toward integration of the field of judgment and decision making.

Reference Notes

1. Adelman, L. Personal communication, 1977.
2. Lonsdale, A. Personal communication, 1977.

References

- Adelman, L. Information about task properties. A necessary, but not a sufficient condition for high levels of achievement in multiple-cue probability learning tasks. (Center for Research on Judgment and Policy Report No. 199). Boulder: Institute of Behavioral Science, University of Colorado, 1977.
- Anderson, N. H. Functional measurement and psychophysical judgment. Psychological Review, 1970, 77(3), 153-170.
- Anderson, N. H. Information integration theory: A brief survey. (Center for Human Information Processing Report No. 24). La Jolla: University of California, San Diego, 1972.
- Anderson, N. H. Information integration theory applied to attitudes about U.S. Presidents. Journal of Educational Psychology, 1973, 64, 1-8.
- Anderson, N. Cognitive algebra. In L. Berkowitz (Eds.), Advances in experimental social psychology (Vol. 7). New York: Academic Press, 1974.
- Anderson, N. H., & Alexander, G. R. Choice test of the averaging hypothesis for information integration. Cognitive Psychology, 1971, 2, 313-324.
- Anderson, N. H., & Shanteau, J. Weak inference with linear models. Psychological Bulletin, 1977, 84, 1155-1170.
- Berlin, Isaiaha. Russian thinkers. New York: Viking Press, 1977.
- Bjorkman, M. Inference behavior in nonmetric ecologies. In L. Rappoport & D. A. Summers (Eds.), Human judgment and social interaction. New York: Holt, Rinehart & Winston, 1973.

References

VII/2

- Bock, R. D., & Jones, L. V. The measurement and prediction of judgment and choice. San Francisco: Holden-Day, 1968.
- Brehmer, B. Cognitive dependence on additive and configural cue-criterion relations. American Journal of Psychology, 1969, 82, 490-503.
- Brehmer, B. Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. Organizational Behavior and Human Performance, 1974, 11, 1-27.
- Brehmer, B. Policy conflict and policy change as a function of task characteristics. IV. The effect of cue intercorrelations. Scandinavian Journal of Psychology, 1975, 16, 85-96.
- Brehmer, B., & Hammond, K. R. Cognitive sources of interpersonal conflict: Analysis of interactions between linear and nonlinear cognitive systems. Organizational Behavior and Human Performance, 1973, 10, 290-313.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: Wiley, 1956.
- Brunswik, E. Wahrnehmung und gegenstandswelt: grundlegung einer psychologie vom gegenstand her (Perception and the world of objects: the foundations of a psychology in terms of objects). Leipzig und Wien: Deuticke, 1934.
- Brunswik, E. The conceptual framework of psychology. In International encyclopedia of unified science (Vol. 1, No. 10). Chicago: University of Chicago Press, 1952.

References

VII/3

- Brunswik, E. Representative design and probabilistic theory in a functional psychology. Psychological Review, 1955, 62, 193-217.
- Brunswik, E. Perception and representative design of psychological experiments (2nd ed.). Berkeley: University of California Press, 1956.
- Brunswik, E. Scope and aspects of the cognitive problem. In H. Gruber, R. Jessor, & K. R. Hammond (Eds.), Contemporary approaches to cognition: A symposium held at the University of Colorado. Cambridge: Harvard University Press, 1957.
- Castellan, N. J. Determination of joint distributions from marginal distributions in dichotomous systems. Psychometrika, 1970, 35(4), 439-454.
- Chapman, L. J., & Chapman, J. P. Genesis of popular but erroneous psychodiagnostic observation. Journal of Abnormal Psychology, 1967, 72, 193-207.
- Chapman, L. J., & Chapman, J. P. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology, 1969, 74, 271-280.
- Cliff, N. Adverbs as multipliers. Psychological Review, 1959, 66(1), 27-44.
- Cochrane, J. L., & Zeleny, M. (Eds.). Multiple criteria decision making. Columbia: University of South Carolina Press, 1973.
- Coombs, C. H., & Bowen, J. Additivity of risk in portfolios. Perception and Psychophysics, 1971, 10, 43-46.

References

VII/4

- Coombs, L. C., Coombs, C. H., & McClelland, G. H. Preference scales for number and sex of children. Population Studies, 1975, 29, 273-298.
- Crow, W. J., & Hammond, K. R. The generality of accuracy and response sets in interpersonal perception. Journal of Abnormal Social Psychology, 1957, 54, 384-390.
- Davidson, D., Suppes, P., & Siegel, S. Decision-making: An experimental approach. Stanford: Stanford University Press, 1957.
- Dawes, R. M. A case study of graduate admissions: Application of three principles of human decision making. American Psychologist, 1971, 26, 180-188.
- Dawes, R. M. The mind, the model, and the task. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. R. Lindman, & D. P. Pisoni (Eds.), Cognitive theory (Vol. 1). Hillsdale, N.J.: Erlbaum, 1975.
- Dawes, R. M., & Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.
- de Neufville, R., & Keeney, R. L. Use of decision analysis in airport development for Mexico City. In A. W. Drake, R. L. Keeney, & P. M. Morse (Eds.), Analysis of public systems. Cambridge, Mass.: M.I.T. Press, 1972.
- Druckman, D. Negotiations. Beverly Hills: Sage Publications, 1977.
- Edwards, W. Probability-preferences in gambling. American Journal of Psychology, 1953, 66, 349-364.
- Edwards, W. The theory of decision making. Psychological Bulletin, 1954, 51, 380-417.

References
VII/5

- Edwards, W. Behavioral decision theory. Annual Review of Psychology, 1961, 12, 473-498.
- Edwards, W. Social utilities. The Engineering Economist, Summer Symposium Series, 1971, 6.
- Edwards, W. Application of research on cognition to man-machine system design. (Engineering Psychology Laboratory Report No. 010342-1-F). Ann Arbor: University of Michigan, 1972.
- Edwards, W. Technology for Director Dubious: Evaluation and decision in public context. In K. R. Hammond (Ed.), Judgment and decision in public policy formation. Boulder: Westview Press, 1978.
- Edwards, W., & Tversky, A. Decision making: Selected readings. Middlesex, England: Penguin Books, Ltd., 1967.
- Einhorn, H. J., & Hogarth, R. M. Unit weighting schemes for decision making. Organizational Behavior and Human Performance, 1975, 13, 171-192.
- Fischer, G. W. Convergent validation of decomposed multi-attribute utility assessment procedures for risky and riskless decisions. Organizational Behavior and Human Performance, 1977, 18, 295-315.
- Fischhoff, B. Attribution theory and judgment under uncertainty. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), New directions in attribution research. Hillsdale, N. J.: Erlbaum Assoc., 1976.
- Fishbein, M., & Ajzen, I. Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, Mass.: Addison-Wesley Publishing Co., 1975.
- Fishburn, P. C. Interdependent preferences on finite sets. Journal of Mathematical Psychology, 1972, 9, 225-236.

References

VII/6

- Frieze, I. H. The role of information processing in making causal attributions for success and failure. In J. S. Carroll & J. W. Payne (Eds.), Cognition and Social Behavior. Hillsdale, N. J.: Erlbaum Assoc., 1976.
- Frieze, I., & Weiner, B. Cue utilization and attributional judgments for success and failure. Journal of Personality, 1971, 39, 591-605.
- Fryback, D. G. Use of radiologists' subjective probability estimates in a medical decision making problem. (Michigan Mathematical Psychology Program Report No. 74-14). Ann Arbor: University of Michigan, 1974.
- Gardiner, P. C. The application of decision technology and Monte Carlo simulation to multiple objective public policy decision making: A case study in California coastal zone management. Unpublished doctoral dissertation, University of Southern California, 1974.
- Gardiner, P. C., & Edwards, W. Public values: Multiattribute-utility measurement for social decision making. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press, 1975.
- Gardiner, P. C., & Ford, A. A merger of simulation and evaluation for applied policy research in social systems. Unpublished manuscript, University of Southern California, 1977.
- Gillis, J. S. The effects of selected anti-psychotic drugs on objective task learning, and interpersonal learning with acute schizophrenics. In K. R. Hammond & C. R. B. Joyce (Eds.), Psychoactive drugs and social judgment: Theory and research. New York: Wiley, 1975.

References
VII/7

- Goldberg, L. R. Man versus model of man: A rationale plus some evidence, for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73(6), 422-432.
- Goldberg, L. R. Man versus model of man: Just how conflicting is that evidence? Organizational Behavior and Human Performance, 1976, 16, 13-22.
- Grether, D. M., & Plott, C. R. Economic theory of choice and the preference reversal phenomenon. American Economic Review, in press.
- Hamilos, C. A., & Pitz, G. F. The encoding and recognition of probabilistic information in a decision task. Organizational Behavior and Human Performance, 1977, 20(2), 184-202.
- Hammond, K. R. Probabilistic functioning and the clinical method. Psychological Review, 1955, 62, 255-262.
- Hammond, K. R. The psychology of Egon Brunswik. New York: Holt, Rinehart, & Winston, 1966.
- Hammond, K. R. Computer graphics as an aid to learning. Science, 1971, 172, 903-908.
- Hammond, K. R. Toward increasing competence of thought in public policy formation. In K. R. Hammond (Ed.), Judgment and decision in public policy formation. Boulder: Westview Press, 1978.
- Hammond, K. R., & Adelman, L. Science, values, and human judgment. Science, 1976, 194, 389-396.
- Hammond, K. R., & Joyce, C. R. B. (Eds.). Psychoactive drugs and social judgment: Theory and research. New York: Wiley, 1975.
- Hammond, K. R., & Kern, F. Teaching comprehensive medical care: A psychological study of a change in medical education. Cambridge, Mass.: Harvard University Press, 1959.

References
VII/8

- Hammond, K. R., Klitz, J. K., & Cook, R. L. How systems analysts can provide more effective assistance to the policy maker. (International Institute for Applied Systems Analysis Research Memorandum RM-77-50). Laxenburg, Austria: IIASA, 1977.
- Hammond, K. R., Mumpower, J. L., & Smith, T. H. Linking environmental models with models of human judgment: A symmetrical decision aid. IEEE Transactions on Systems, Man, and Cybernetics, 1977, SMC-7(5), 358-367.
- Hammond, K. R., Rohrbaugh, J., Mumpower, J., & Adelman, L. Social judgment theory: Applications in policy formation. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1977.
- Hammond, K. R., Stewart, T. R., Adelman, L., & Wascoe, N. E. Report to the Denver city council and mayor regarding the choice of handgun ammunition for the Denver police department. (Center for Research on Judgment and Policy Report No. 179). Boulder: Institute of Behavioral Science, University of Colorado, 1975.
- Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. Social judgment theory. In M. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes: Formal and mathematical approaches. New York: Academic Press, 1975.
- Harsanyi, J. C. Nonlinear social welfare functions. Theory and Decision, 1974, 7, 61-82.
- Harvey, J. H., Ickes, W. J., & Kidd, F. R. (Eds.). New directions in attribution research. Hillsdale, N.J.: Erlbaum Assoc., 1976.

References
VII/9

- Hawkins, D. The science and ethics of equality. New York: Basic Books, 1977.
- Heider, F. Ding und Medium. Symposium, 1926, 1, 109-157.
- Heider, F. The psychology of interpersonal relations. New York: Wiley, 1958.
- Holton, G. Thematic origins of scientific thought: Kepler and Einstein. Cambridge, Mass.: Harvard University Press, 1973.
- Jones, E. E., & Davis, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 2). New York: Academic Press, 1965.
- Jones, E. E., Davis, K. E., & Gergen, K. J. Role playing variations and their informational value for person perception. Journal of Abnormal and Social Psychology, 1961, 63, 302-310.
- Jones, E. E., & Goethals, G. R. Order effects in impression formation: Attribution context and the nature of the entity. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. Morristown, N.J.: General Learning Press, 1972.
- Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. Attribution: Perceiving the causes of behavior. Morristown, N.J.: General Learning Press, 1972.
- Jones, E. E., & McGillis, D. New directions in attribution theory research. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), New directions in attribution research. Hillsdale, N.J.: Erlbaum Assoc., 1976.

References

VII/10

- Kahneman, D., & Tversky, A. Subjective probability: A judgment of representativeness. Cognitive Psychology, 1972, 3, 430-454.
- Kahneman, D., & Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80(4), 237-251.
- Kahneman, D., & Tversky, A. Prospect theory: An analysis of decision under risk. Jerusalem: Hebrew University, 1977.
- Kahneman, D., & Tversky, A. Intuitive prediction: Biases and corrective procedures. Management Science, in press.
- Kaplan, M. F., & Schwartz, S. (Eds.). Human judgment and decision processes. New York: Academic Press, 1975.
- Keeney, R. L. A decision analysis with multiple objectives: The Mexico City airport. Bell Journal of Economics and Management Science, 1973, 4, 101-117. (a)
- Keeney, R. L. A utility function for the response times of engines and ladders to fires. Urban Analysis, 1973, 1, 209-222. (b)
- Keeney, R. L. Preference models of environmental impact. (International Institute of Applied Systems Analysis Research Memorandum RM-76-4). Laxenburg, Austria: IIASA, 1976.
- Keeney, R. L. The art of assessing multiattribute utility functions. Organizational Behavior and Human Performance, 1977, 19, 267-310.
- Keeney, R. L., & Raiffa, H. Decisions with multiple objectives: Preferences and value tradeoffs. New York: Wiley, 1976.
- Keeney, R. L., & Sicherman, A. An interactive computer program for assessing and analyzing preferences concerning multiple objectives. Behavioral Science, 1976, 21, 173-182.

References
VII/11

- Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), Nebraska symposium on motivation, 1967 (Vol. 15). Lincoln: University of Nebraska Press, 1967.
- Kelley, H. H. Causal schemata and the attribution process. New York: General Learning Press, 1971.
- Kelley, H. H. The processes of causal attribution. American Psychologist, 1973, 28, 107-128.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. Foundations of measurement (Vol. 1). New York: Academic Press, 1971.
- Krantz, D. H., & Tversky, A. Conjoint-Measurement analysis of composition rules in psychology. Psychological Review, 1971, 78(2), 151-169.
- Krech, D. Discussion: Theory and reductionism. Psychological Review, 1955, 62(3), 229-230.
- Kukla, A. Foundations of an attributional theory of performance. Psychological Review, 1972, 79, 454-470.
- Lampel, A. K., & Anderson, N. H. Combining visual and verbal information in an impression-formation task. Journal of Personality and Social Psychology, 1968, 9, 1-6.
- Leal, L., & Pearl, J. An interactive program for conversational elicitation of decision structures. IEEE Transactions on Systems, Man, and Cybernetics, 1977, SMC-7(5), 368-376.
- Lichtenstein, S. C., & Slovic, P. Reversals of preference between bids and choices in gambling decisions. Journal of Experimental Psychology, 1971, 89, 46-55.

References
VII/12

- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, L., & Combs, B.
Perceived frequency of lethal events. (Decision Research Report
76-2). Eugene, Ore.: Decision Research, 1978.
- Lindell, M. K., & Stewart, T. R. The effects of redundancy in
multiple-cue probability learning. American Journal of Psychology,
1974, 87, 393-398.
- Luce, R. D. Individual choice behavior. New York: Wiley, 1959.
- Luce, R. D., & Raiffa, H. Games and decisions. New York: Wiley, 1957.
- Mosteller, F., & Nogee, P. An experimental measurement of utility.
Journal of Political Economics, 1951, 59(5), 371-404.
- Mumpower, J. A comparison of attribution, social learning, and social
judgment theories. (Center for Research on Judgment and Policy
Report No. 208). Boulder: Institute of Behavioral Science,
University of Colorado, 1976.
- Mumpower, J. L., & Hammond, K. R. Entangled task-dimensions: An
impediment to interpersonal learning. Organizational Behavior and
Human Performance, 1974, 11, 377-389.
- McArthur, L. A. The how and what of why: Some determinants and
consequences of causal attribution. Journal of Personality and
Social Psychology, 1972, 22, 171-193.
- McClelland, G. H. Psychological processing of evaluative meaning of
adverb-adjective combinations. (Michigan Mathematical Psychology
Program Technical Report 1974-11). Ann Arbor: University of
Michigan, 1974.

- McClelland, G. H. Equal versus differential weighting for multiattribute decisions: There are no free lunches. (Center for Research on Judgment and Policy Report No. 207). Boulder: Institute of Behavioral Science, University of Colorado, 1978.
- McClelland, G., & Rohrbaugh, J. Who accepts the Pareto axiom? The role of utility and equity in arbitration decisions. (Center for Research on Judgment and Policy Report No. 202). Boulder: Institute of Behavioral Science, University of Colorado, 1977.
- McClelland, L., & Auslander, N. Perceptions of crowding and pleasantness in public settings. Environment and Behavior, in press.
- McClintock, C. G. Social motivation in settings of outcome interdependence. In D. Druckman (Ed.), Negotiations: Social-psychological perspectives. Beverly Hills: Sage Publications, 1977.
- Nisbett, R. E. & Borgida, E. Attribution and the psychology of prediction. Journal of Personality and Social Psychology, 1975, 32, 932-943.
- Nisbett, R. E., & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84(3), 231-259.
- Ostrom, T. M., Werner, C., Saks, M. J. An integration theory analysis of juror's presumptions. Journal of Personality and Social Psychology, in press.
- Pepper, Stephen. World hypotheses. Berkeley: University of California, 1948.
- Phelps, R. H., & Shanteau, J. Livestock judges: How much information can an expert use? Organizational Behavior and Human Performance, in press.

References
VII/14

- Phillips, L., & Edwards, W. Conservatism in a simple probability inference task. Journal of Experimental Psychology, 1966, 72(3), 346-354.
- Raiffa, H. Decision analysis: Introductory lectures on choices under uncertainty. Reading, Mass.: Addison-Wesley, 1968.
- Raiffa, H. Preferences for multi-attributed alternatives. (RM-5868-DOT/RC). Santa Monica: The Rand Corporation, 1969.
- Rapoport, A. (Ed.). Game theory as a theory of conflict resolution. Dordrecht, Holland: D. Reidel Publishing Co., 1974.
- Rappoport, L., & Summer, D. A. (Eds.). Human judgment and social interaction. New York: Holt, Rinehart & Winston, 1973.
- Rawls, J. A theory of justice. Cambridge, Mass.: Harvard University Press, 1971.
- Roche, J. G. Preference tradeoffs among instructional programs: An investigation of cost-benefit and decision analysis techniques in local educational decision making. Unpublished doctoral dissertation, Graduate School of Business Administration, Harvard University, 1971.
- Rohrbaugh, J. Cognitive maps: Describing the policy ecology of a community. Great Plains-Rocky Mountain Geographical Journal, 1977, 6, 64-73.
- Ross, L. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 10). New York: Academic Press, 1977.

- Schmidt, F. L. The relative efficiency of regression and simple unit predictor weights in applied differential psychology. Educational and Psychological Measurement, 1971, 31, 699-714.
- Shanteau, J. C. An additive model for sequential decision making. Journal of Experimental Psychology, 1970, 85, 181-191.
- Shanteau, J. Descriptive versus normative models of sequential inference judgment. Journal of Experimental Psychology, 1972, 93, 63-68.
- Shanteau, J. An information-integration analysis of risky decision making. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press, 1975.
- Shanteau, J. C., & Anderson, N. H. Test of a conflict model for preference judgment. Journal of Mathematical Psychology, 1969, 6, 312-325.
- Shanteau, J., & Nagy, G. Decisions made about other people: A human judgment analysis of dating choice. In J. Carroll & J. Payne (Eds.), Cognitive and social judgment. Potomac, Md.: Erlbaum Assoc., 1976.
- Shanteau, J., & Phelps, R. H. Judgment and Swine: Approaches and issues in applied judgment analysis. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1977.
- Shaver, K. An introduction to attribution processes. Cambridge, Mass.: Winthrop, 1975.
- Shephard, R. N. Use of judgments in making optional decisions. Paper presented at the meeting of the American Association for the Advancement of Science, 1962.

References
VII/16

- Slovic, P. Towards understanding and improving decisions. In E. I. Salkovitz (Ed.), Science, technology, and the modern Navy: Thirtieth anniversary 1946-1976. Arlington, Va.: Department of the Navy, Office of Naval Research, 1976.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. In Annual Review of Psychology, 1977, 28, 1-39.
- Slovic, P., & Lichtenstein, S. C. The importance of variance preferences in gambling decisions. Journal of Experimental Psychology, 1968, 78, 646-654.
- Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.
- Starr, M. K., & Zeleny, M. MCDM--State and future of the arts. TIMS Studies in the Management Sciences, 1977, 6, 5-29.
- Thrall, R. M., Coombs, C. H., & Davis, R. L. (Eds.). Decision processes. New York: Wiley, 1954.
- Timmers, H., & Wagenaar, W. A. Inverse statistics and misperception of exponential growth. Perception and Psychophysics, 1977, 21, 558-562.
- Tversky, A. Additivity, utility and subjective probability. Journal of Mathematical Psychology, 1967, 4(2), 175-202.
- Tversky, A. Intransitivity of preferences. Psychological Review, 1969, 76, 31-48.
- Tversky, A. Elimination by aspects: A theory of choice. Psychological Review, 1972, 79(4), 281-299. (a)

References
VII/17

- Tversky, A. Choice by elimination. Journal of Mathematical Psychology, 1972, 9(4), 341-367. (b)
- Tversky, A. Features of similarity. Psychological Review, 1977, 84(4), 327-352.
- Tversky, A., & Kahneman, D. Belief in the law of small numbers. Psychological Bulletin, 1971, 76(2), 105-110.
- Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 1973, 5, 207-232.
- Tversky, A., & Kahneman, D. Causal schemata in judgments under certainty. In M. Fishbein (Ed.), Progress in social psychology. Hillsdale, N.J.: Erlbaum Assoc., 1977.
- Urdang, L. (Ed.). Random House dictionary (college edition). New York: Random House, 1968.
- von Neumann, J., & Morgenstern, O. Theory of games and economic behavior (2nd ed.). Princeton, N.J.: Princeton University, 1947.
- Wainer, H. Estimating coefficients in linear models: It don't make no nevermind. Psychological Bulletin, 1976, 83, 213-217.
- Weiner, B., & Kukla, A. An attributional analysis of achievement motivation. Journal of Personality and Social Psychology, 1970, 15, 1-20.
- Wortman, C. Causal attributions and personal control. In J. Harvey, J. Ickes, & R. Kidd (Eds.), New directions in attribution research. Hillsdale, N.J.: Erlbaum Assoc., 1976.
- Wyer, R. S. Information redundancy, inconsistency, and novelty and their role in impression formation. Journal of Experimental Social Psychology, 1970, 6, 111-127.

References

VII/18

Zeleny, M. On the inadequacy of the regression paradigm used in the study of human judgment. Theory and Decision, 1976, 7, 57-65.

Zeleny, M. Adaptive displacement of preferences in decision making. TIMS Studies in the Management Sciences, 1977, 6, 147-157.

A P P E N D I X A

COMMENT BY

RALPH KEENEY

WOODWARD-CLYDE CONSULTANTS

Ralph Keeney

My comments will be divided into two parts. The first part concerns my perceptions of decision theory--its theory, its method, and its procedures, and the second part concerns a few overall comments on the report. The discussion of the theory will be categorized as in the report. I'll discuss its origins, scope, intended functions, principal concepts, loci of concepts, and intended uses. The method and procedures sections will also be discussed to be similar to the presentation in the Colorado report.

THEORY

Two key references which are not included here concerning the origins of decision theory are Ramsey and Savage. An important point is the following: although decision theory does have much of its basis and origins in economic theory with what is referred to as the "rational man" theory, I believe the current focus is a little bit different. In particular, I consider decision theory to take a prescriptive approach rather than a normative approach. The prescriptive approach takes as basic the values and judgments of the person or the entity making the decision and the logic which that person or entity would like to use on the decision. From this it tries to infer what actions should be taken. My understanding of the rational man procedure is that someone else would prescribe the logic which they would define to be rational such that "any rational man should use." The main point is that in a prescriptive theory, the assumptions of rationality should be verified with the decision-making entity to decide if

in fact the proposed method and procedures are appropriate for him or her or it in the case of an entity.

Concerning the scope of decision theory I think the limitation to the single system case is not exactly appropriate. Certainly, most of the decision theory analysis to date do concern this case. I think that is more a limitation of some of the theory that has been developed than it is a limitation to the logic which the theory would use. Some of the decision theory work on social welfare functions does, as mentioned, convert the problem into a single system case where the overall decision maker assigns the value tradeoffs between the individuals who are members of the group. However certainly those people who have contributed to this area have made it clear that if in fact it is an n-system case, then the group of individuals, that is the n, must somehow collectively decide upon what the value tradeoffs are between themselves.

Concerning the intended function of decision theory, the report states that a criterion for the validity of the theory is its logical, mathematical consistency. I believe other criteria are the acceptance of the insights and sometimes the implications of the analyses conducted by the decision makers. Throughout this section on intended function of the theory, it is pointed out that the purpose is something like to achieve logical consistency and find out which decision should be taken. I believe that many of the reasons for conducting the analysis are actually stated at the end of the theory section of the Colorado report. Specifically, I refer to the addendum on intended uses of the results. The theory is supposed to aid clear thinking. It is supposed to educate the decision maker. It is supposed to facilitate communication, etc.

The report states that decision theory is not concerned about why decision makers deviate from the logic of decision theory. I do not think this is the

case. Certainly an interesting recent case by Hax and Wiig indicates an example where in fact the decision maker did not believe the analysis. On further investigation in this case it was found that the reason was that some important objectives of the decision maker were not formally included in the analysis. Once these were included, the decision maker felt more comfortable with the implications of the study and felt that he had gained some additional insights useful for helping him make the decision.

On page II/D/3 it is stated that decision theory considers its axioms to be reasonable and desirable rules for decision making behavior that everyone would want to follow, once they understood them. This is not what I believe. These axioms are reasonable for me. I also believe that in many cases, an individual with a decision to make may feel that they are reasonable for them for that decision. If this is the case then the implications of the axioms are reasonable to use in prescribing what the decision maker should in fact do for this particular problem.

Many places in this report, such as II/D/18, it is stated that the results of PDT research challenge the validity and the premises of decision theory. I do not feel this statement really captures all of what is implied by those results. It certainly implies that descriptively many times the SEU approach is not appropriate to describe behavior and it may imply that this approach is not appropriate to prescribe behavior in some or many situations. However, if an individual says for a particular problem that they want the axioms of decision theory to be followed, then one implication of the PDT research is that it is best to do this formally, simply because individuals are not very good at processing and integrating the relevant information in their mind. Conducting a formal analysis at least gives us a chance to catch the shortcomings. And the

PDT research indicates particular ones to guard against.

Throughout section E there are also many comments concerning the challenges to decision theory from PDT. I view many of the results there not as challenges but as a collective body of very relevant material to one who would purport to do decision analysis of problems. In particular I agree very strongly with the statement of Paul Slovic beginning of page 23 of that section. "Subjective judgments of probability and utility are essential inputs to decision analysis. We still do not know the best ways to elicit these judgments. Now that we understand many of the biases to which these judgments are susceptible, we need to develop debiasing techniques to minimize their destructive effects." A good decision theorist would certainly keep informed of this work and both try to adapt his or her practice to account for them and try to avoid falling into the obvious traps pointed out by this research.

Concerning the intended use of the theory, again I think there is an oversimplification where it states that in decision theory there is no theoretical attention given to the process by which past and present information or data is transformed into subjective probabilities. In many cases models are developed which do try to discern exactly what bits of information are being integrated in order to arrive at a subjective probability. The analogous process certainly takes place when one has multiattribute utility functions. One divides the problems into the various components and formally indicates how they are integrated together to get the overall utility of a particular consequence. Thus, attention is given to exactly how bits of information are combined into overall information. There is always the question of how far back one should go, but in any particular problem the judgment of the decision analyst and of the decision maker or of the decision makers must be exercised in this choice.

With regard to the intended uses of decision theory, I do not feel that better is defined in terms of the approximation of the procedure to the axioms of decision theory. Better has a much broader definition, and in some sense, better is defined as whether the decision maker feels it was better to use decision theory than it was to use the other alternatives he or she had available for making their decision. The logical consistency of the mathematics is not all that is needed to support decision theory. To me it is crucial that the decision maker understand the axioms and accept them for the theory to be relevant to them at all.

To conclude this section I would like to say that I do feel, used well, decision theory does lead to those claims discussed in the addendum to section G. From my point of view I feel decision theory was represented in this section as being a little bit narrower, or should I say narrow-minded, than in fact I would like to think it is. Secondly, I do not feel that many of the results in behavioral decision theory and psychological decision theory are contrary to what is being done in decision theory. In particular, this is true of the applications. There has been and is a lot to be learned from the research and results of those areas that decision analysts should be aware of and should bring to bear in their practice. Perhaps collectively we are not as good at doing that as we should be, but I believe the overall implications of the research in these areas support each other much more than it does conflict.

METHOD

I will now discuss decision theory in terms of each of the five dimensions outlined in the Colorado report. In nomothetic analysis, there may be many cases where the judges informally use an idiographic first step. That is, each judge

combines many things in his or her own mind before giving their judgment about that aspect. A nomothetic step occurs often in decision theory when it is required that experts judgments or the preferences of different people over different attributes be combined. To me, the critical issue in how one should proceed is partly indicated by the specific problem being addressed.

In the section on sampling domains, I agree there is usually one decision maker in decision theory. However, the statement that the emphasis is on collecting as much information from that decision maker as is necessary to construct the required probability distributions and utility functions is not completely correct in my mind. I feel one wants to collect enough information to understand and structure the problem, to conduct an analysis of the various options, and to gain some insights and help for the decision maker who must make the final decision. The distinction is one which shows up throughout the report; I feel good decision theory is problem oriented and not technique oriented.

With regard to stimulus-objective composition, it is true that one often uses hypothetical and implausible profiles describing an object. However this is certainly not a limitation of the theory, and with the computer backup that is now developed, such limitations can certainly be removed.

With regard to judgment decomposition, certainly one tries to verify a set of independence conditions at the beginning of the problem. However, if the independence conditions are not met, and in particular if they are seriously violated, one does not browbeat the decision maker into accepting the assumptions. Rather one tries to identify the dependencies and the reasons for them, and then either build a model that incorporates those dependencies or alter the structure of the problem by adding another attribute, for instance, such that the dependencies are no longer crucial.

With regard to partitioning, early in this section it implies that knowledge has entered decision theory usually only in the form of subjective probabilities. Later I think it is stated that this is not the case. Specifically, models, both simulation and analytic, as well as data, subjective judgments and results of experiments are used to quantify individual's knowledge.

PROCEDURES

Basically I felt the procedures outlined concerning decision theory were quite accurate. This is especially true when one bears in mind the caveats at the beginning of the chapter stating that the procedure used depends on among other things the problem being addressed. There are a couple of minor points I'd like to make about the utility assessment section, however. First, on page 63, the two-attribute multilinear form is certainly not the most general utility function with two attributes. On page 70, it is stated that utility functions are usually assessed first and then the weights are assessed in decision theory. In many problems I feel this situation is reversed.

Concerning a more important point, on page 69 it is stated that several researchers have demonstrated that in many instances all possible sets of weights produce essentially the same results. The important point here is that the results referred to concern the evaluation of the alternatives. If one broadens the results to include gaining a better structure and understanding of the problem, a more appropriate set of attributes, and more insights from this more detailed assessment of the value structure, then I believe the results are not the same.

GENERAL COMMENTS

I have five separate general comments which may be of relevance.

- (1) The first concerns the question of whether integration of approaches

is desirable or useful. For this I feel one needs to ask the question, why should one want to integrate? For instance, if we integrate the approaches to improve decision making, the integration may be different than if we were doing it to improve our understanding of how decisions are made.

(2) I feel that a hierarchical integration, if integration is done, may be a more useful way to address the problem. That is, I would maybe integrate the three approaches categorized as group one together, and similarly integrate the three approaches categorized as group two. Then I would integrate these two groups together.

(3) Throughout it was indicated several times that behavioral decision theory and psychological decision theory were at odds with decision theory. I feel that this is not really the case, and that what is learned in each of these three approaches is useful and relevant to what is done and what should be done in the other approaches. The results in one approach are relevant for both the research and the applications within the others. I view them much more as complimentary bodies of knowledge than contradictory or conflicting bodies of knowledge.

(4) I think a better indication of other groups who may be considered in the decision theory category would be worthwhile. In particular I think those people who have been affiliated with the Stanford Research Institute group, Decisions and Designs, Decision Focus, Applied Decision Analysis, etc., should probably be identified. Furthermore, their methods and procedures may be somewhat different than the ones commonly categorized throughout this report as the decision theory approach.

(5) I feel some of the inferences drawn from a footnote, an item included in an index, or a quotation out of context were stronger than sometimes warranted.

AD-A061 750

COLORADO UNIV BOULDER CENTER FOR RESEARCH ON JUDGME--ETC F/G 5/10
THE COLORADO REPORT ON THE INTEGRATION OF APPROACHES TO JUDGMEN--ETC(U)
OCT 78 K R HAMMOND, G H MCCLELLAND

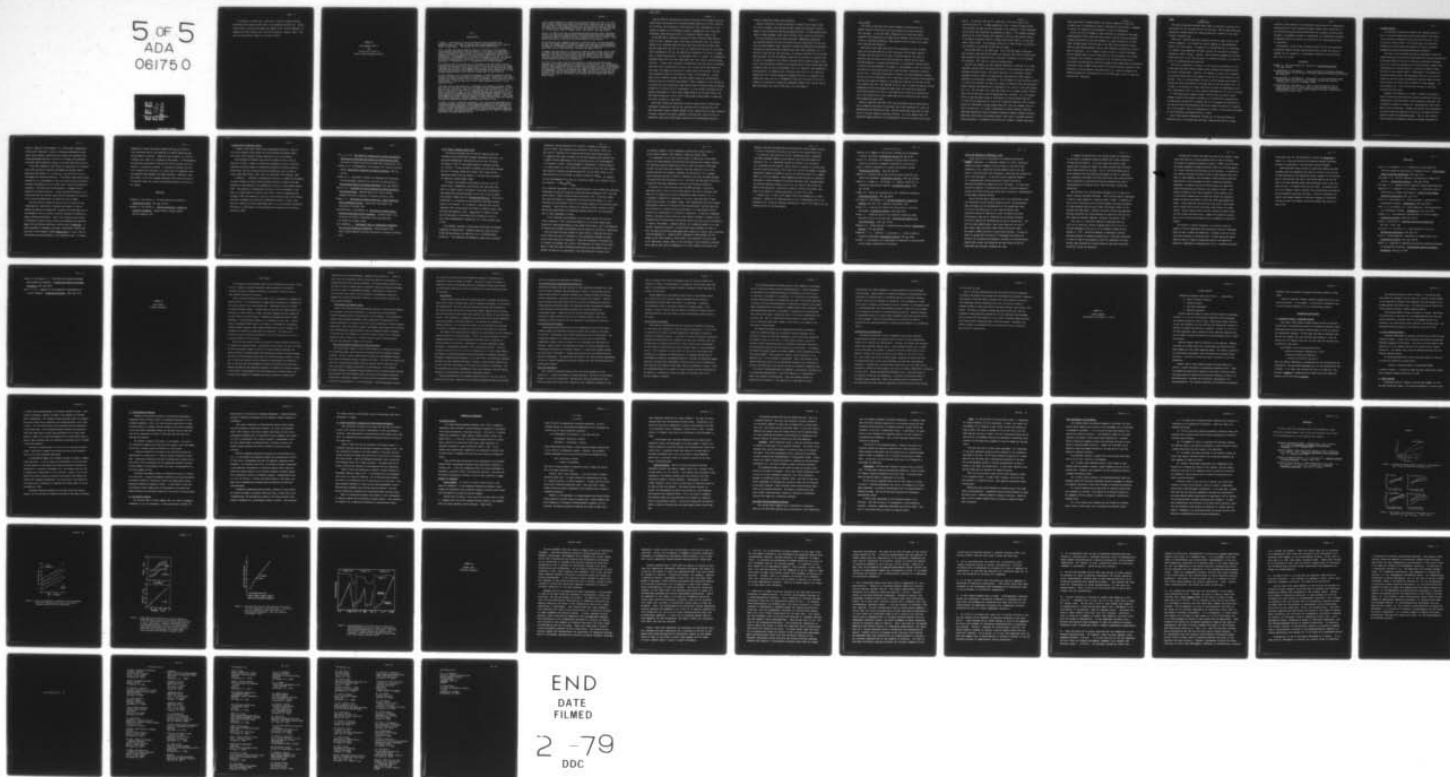
N00014-77-C-0336

UNCLASSIFIED

CRJP-213

NL

5 OF 5
ADA
06175 0



In conclusion, I would like to state that I certainly learned something from reading this report and felt that it was worthwhile and done well. Because we were asked to criticize the report and comment on how it may be improved, my comments have been limited to this side of the question. However, again I feel there are many positive aspects to the overall report.

COMMENT BY
WARD EDWARDS (PART I)
AND
RICHARD JOHN (PART II)
SOCIAL SCIENCE RESEARCH INSTITUTE

Part I

(Ward Edwards)

I wonder if your book may not be an appropriate place to question the appropriateness of the idea of sampling used in the metaphorical sense that is normal to psychology. The reasons why will become clearer below.

First, consider the idea of sampling subjects. To the best of my knowledge, no experimental psychologist has ever done so in the history of psychology. A sample is defined as a subset chosen by a suitably representative method from a population. But what population do psychologists sample from, and how representative is the method? We in fact use as Ss whoever we can get, our methods of selection have virtually nothing to do with representativeness, and the population, if definable at all, is the population of college students.

This tradition did little harm when it originated in the psychology of sensory processes (and perhaps also did little harm in the early work on learning, though I am less sure of that). The argument could be, was, and is made that the sample represents the population with respect to the process being studied--that is, that college students differ little from the village idiot in visual ability.

But what nonsense to try to extend that argument to cognitive processes! College students are chosen for excellence at cognitive processes. Research on other populations would unquestionably lead to inferior performance. Would there be structural similarities? I have no idea, and I am no longer prepared to accept it as an article of faith.

This bears also on the discussion of idiographic and nomothetic analyses. Almost never are we interested in the individual whom we study idiographically; we simply hope (on what basis?) that others are sufficiently like him for it to be worthwhile to study what he does. As for nomothetic analyses that average over Ss, this simply means that we generalize from n non-randomly-selected members of an undefined population to that undefined population, which seems no better. I don't see the argument that makes n college students all doing the same things more representative of the generalized normal adult human mind than only one.

In short, at least in cognitive psychology, I don't see how an experiment can be considered anything other than a set of carefully managed clinical observations. I value the careful management--but the generalization process is completely intuitive, and the idea that we are sampling in a manner that represents cognitive processes across a defined population of individuals (and of tasks) is completely misleading. You might usefully say so.

I have related objections to the notion of sampling objects and tasks. All that really means is that we use several different ones, rather than one. We may also have some idea of what these differences are. But we certainly have no idea what the population is, and consequently no basis for generalizing. All we really can say is that several kinds of objects and/or tasks lead to the same effects.

Finally, we come to the idea of representative design. Brunswik made a meaningful effort to be genuinely representative; he followed people around and asked them to judge how long a stick was in various situations. While this is crude, it does embody some notion of what might be called sampling of the ecology.

But we certainly don't manage to do that in cognitive tasks. We have absolutely no idea, for example, what are cognitive tasks, much less what is the ecological distribution of incidence of them. (For whom?) Nor do we know anything about the population of cue intercorrelations. So how can we sample from it?

It seems to me that, here as in the case of sampling across people, we are simply using the idea that several are better than one, especially if the several (of whatever it is) lead to similar results. But why dignify that with a name that has a well-defined statistical meaning? And why not acknowledge the extremely limited nature of our ability to generalize?

This whole issue seems to me far more important in cognitive psychology than anywhere else. We are attempting to study one of the most human, i.e. idiosyncratic kinds of human performance. Either we should define carefully a domain of performance and then really sample it (which so far as I know we almost never do, unless "domain of performance" is taken as synonymous with domain of laboratory tasks), or else we should admit that we study specific tasks that interest us, pick the most interesting ones, explain why they are most interesting, and forget about further generalization.

Decision theorists and behavioral decision theorists alike recognize that one cannot assess the probability of anything without specifying it first; similarly for utilities. The raw material of both points of view, consequently, is not vectors of numbers to be integrated by means of aggregation rules, but rather structures--often structures similar in nature and intent to those of SJT.

DT and BDT tend to represent structures as trees. There are three familiar kinds of trees: inference trees, value trees, and decision trees. For reasons obscure to us, inference trees and value trees grow from top to bottom, while decision trees grow from left to right. An inference tree has data at its bottom level, hypotheses at the top, and subordinate hypotheses at all levels in between. Thus, in photo-reconnaissance, the photograph itself is a datum. An intermediate hypothesis might be of the form "Those blobs look like a tank column moving down the road". A higher level hypothesis incorporating other data would be "Yesterday there was a tank concentration at point A, well to the rear of the location of the column. The enemy must be moving his tanks closer to the action". The top level hypothesis might be "The enemy will attack in force tomorrow." At each step of the inferential process, more data become relevant to the hypothesis; different chains of inference come together. Formal treatments of such inferential structures are entirely possible; they exhibit either extremely strong and numerous assumptions of various kinds of independence, or else extremely elaborate and demanding requirements for conditioning judged probabilities on other things. The most sophisticated work in this field is being done by David Schum, in analysing the formal structure of legal proof.

Value trees reflect the hierarchical nature of human values. A school board interested in educational quality, may (and did) wish to subdivide that into such lower-order values as teacher-student ratios, contact hours per week, learning of basics, provision for special programs of various kinds, and so on (for 13 dimensions, some further subdivided). Realistic utility measurement would be

virtually impossible without such structures.

Decision trees have a curious oscillatory structure: an act leads to one of a set of outcomes. Each possible outcome creates the situation in which a number of new alternatives acts may be appropriate. Each of these possibilities lead to further outcomes, which lead to further acts, and so on to infinity.

All tree structures can and should be pruned. Independence assumptions and ignoring branches with very low probability and unspectacular utility both help to prune inference trees. Ignoring of unimportant dimensions and resistance to excessive subdivision help to prune value trees. Sheer ignorance, i.e. inability to look very far into the future, is the most important tool for pruning decision trees. The utility attached to each twig at the end of a decision tree is really a surrogate for the infinite structure that has been cut away at that point. The difficulties of defining or conceptualizing ultimate values baffle DTs and BDTs as much as they do philosophers. But it is important, both conceptually and practically, to recognize that (virtually) every utility is an intuitive compression of an infinite sequence of expected utilities lying beyond it but pruned; thus every utility is fundamentally risky and the notion of utility is primarily a surrogate for unknown and unknowable probabilities. (Not all DTs or BDTs would agree; the views of PDTs about this are unknown.)

P II - D 4-9:

The notions of BDT about the relation between the prescriptions for optimal behavior offered by DT and actual behavior have been developing over the years. In the late 1960's, Edwards and his associates, mostly concerned with conservatism in probabilistic inference, were willing to think of men as conservative Bayesians. They tended to think of them also as somewhat inaccurate SEU maximizers.

These basic views remain relatively firmly entrenched in their thinking today, but have been modified to take into account the extremely powerful influence that the nature of the task, elicitation methods, and state of education used can have on the resulting behavior. Whether men are rather accurate Bayesians, or conservative Bayesians, or non-Bayesian, depends heavily on exactly how the experiment is designed. It also depends on level of training, availability of aids (both outside and inside the head), experience, expertise, and the like. Traditional psychological reliance on college sophomores; complex, boring, and trivial experiments; and in some cases casual-to-uncontrolled procedures combined with compelling examples on the page have combined to lead some non-DTs to an image of man that, from some BDT points of view, is highly misleading. College sophomores working with bookbags and poker chips are conservative Bayesians; intelligence analysts are less so; radiologists are virtually not conservative at all when dealing with x-rays--though they too are conservative when dealing with bookbags and poker chips.

Recently, Mosteller and Tukey (1977) have been emphasizing the importance of using systematic exploration of deviations from linear regression as a basis for understanding the lawfulness of data. BDT has consistently taken much the same point of view with regard to deviations from DT. At a crude enough level, the standard DT models provide a first approximation (often an excellent one) to what

men do. In any case, they are in a sense part of the task as well as of what men bring to it. It seems appropriate, then, to remove DT-based effects from the data, and then look for lawfulness in the residuals. If that lawfulness can be found, then psychology has something to explain. This research strategy underlies all of what Edwards calls cognitive illusions, from the halo effect in syllogistic reasoning (1930's), through conservatism (1960's), to the judgmental biases of PDT. This strategy, along with a lingering hope of explaining such residuals, are what BDT and PDT have in common.

BDT and PDT differ in two major ways. One is that BDT seems to take much more seriously than does PDT the effects on residuals left after DT effects have been removed of task characteristics, individual differences, prior knowledge and expertise. This class of issue was discussed above.

The other and probably far more important difference is that BDT (or at least some BDTs) and PDT differ both about the nature of the psychological enterprise and about the nature of man. The notion that man is an intellectual cripple presented by Slovic (1976), and supported by a good many other PDTs, seems to some BDTs to be absolutely preposterous. Slovic manages to write excellent articles in spite of his handicap, and some of us manage to understand them. Men discovered Bayes's Theorem, the maximization of expected utility, and all of the other apparatus of DT. Men use these tools, when the stakes are high enough to make it worthwhile and the individuals know and appreciate their usefulness. It is as inappropriate to describe human beings as intellectually crippled because they don't all know and use these tools as it would be to describe those who visit a foreign country as intellectual cripples because they don't know the language. They can learn, and may do so. Is a person intellectually crippled because he does not understand enumeration? Simple arithmetic? Complex arithmetic? Algebra? Calculus? Where does nature leave off and nurture begin? What level of equipment acquired during growing up is considered to be normal for Titchener's "generalized normal

adult human mind?" How much external and internal apparatus is that mind allowed to use in performing its tasks? Psychologists have scarcely attempted to answer this question--which admittedly is close to unanswerable.

- Facing this kind of dilemma, at least some BDTs are prepared to soft-pedal sweeping generalization about the nature of man. Some men are exact Bayesians (perhaps because they can do the arithmetic on a desk calculator or in their heads; perhaps because intuition does it for them, given appropriate tasks); some are approximate Bayesians; some are poor Bayesians--and there are various ways of being poor. It all depends--on man, on task, on state of knowledge. Similar comments apply to use of other parts of the DT apparatus. This change of viewpoint, which amounts to rejection of the Titchenerian definition of the purpose of psychology that has been accepted by most experimental psychologists since then, biases such BDTs toward study of elicitation methods, methods of task formulation, training methods, and applications--i.e. toward the numerous and effective methods for invalidating Slovic's and PDT's image of man, at least outside the PDT's laboratory.

The scope of Behavioral Decision Theory (BDT), as described in section II/C, does justice to only the empirical half of the story. BDT has made significant theoretical advances beyond the "single-system case," primarily on the topic of model (in)sensitivity.

Within the "double-system case" paradigm outlined on pages II/C/2-3, subjective "judgments" of a subject are usually compared to some mathematical model of the "task system". The degree of fit between the two is interpreted as a measure of how adequately the model accounts for the cognitive processes presumed to control the S's response. When subjective judgments do not match model predictions, we say that the model is inaccurate and hence, not useful as a tool for understanding and prediction.

The implicit assumption, of course, is that subjective responses are valid, in the sense that they reflect something about inaccessible processes carried on inside a subject's head. Typically, one attempts to justify this assumption by providing instructions to the subject in the form of a payoff matrix or loss function. That is, the experimenter defines a correct response and proceeds to arrange the reward contingencies of the experiment so that the subject will benefit most if he responds in a manner consistent with what the experimenter has in mind. (See Edwards, 1961, for a more complete development of this line of thought)

An intriguing question emerges however: How efficient are typical payoff schemes at getting the subject to respond so as to maximize his expected gain? In other words, how much will a subject lose if his responses are insensitive to the implicit instructions provided by the payoff scheme? The question is answered for a fairly general class of circumstances in a series of three unpublished papers by von Winterfeldt and Edwards (1973a, b, 1975).

A few of the important theoretical findings are: 1) decision-theoretical maxima are flat, in the sense that seriously inappropriate behavior produces

relatively little reduction in the subjects' expected payoff; 2) mathematical characteristics of linear optimization models impose severe restrictions on the functions evaluating choice alternatives such as gambles, multiattribute outcomes, or consumption streams; 3) nevertheless, by inadvertently selecting a dominated decision alternative, the possibility for large expected losses is substantial.

Unfortunately, little of the information gained from the von Winterfeldt and Edwards research has been applied outside BDT to either real decision problems (DT, for example) or in experimental research on decision making (PDT, SJT, IIT, or AT).

REFERENCES

- Edwards, W. Costs and Payoffs are instructions. Psychological Review, 1961, 68, 275-284.
- von Winterfeldt, D. and Edwards, W. Costs and Payoffs in Perceptual Research. Technical Report, The University of Michigan, Engineering Psychology Laboratory No. 011313-1-T
- von Winterfeldt, D. and Edwards W. Flat maxima in linear optimization models. Technical Report, The University of Michigan, Engineering Psychology Laboratory, No. 011313-4-T, November, 1973b.
- von Winterfeldt, D. and Edwards, W. Error in Decision Analysis: How to Create the Possibility of Large Losses by Using Dominated Strategies. Research Report, 75-4, Social Science Research Institute, the University of Southern California, 1975.

Intended Function

As correctly pointed out by Hammond, the intended function of BDT is to discover departures from the prescriptions of DT and to invent corrective procedures to deal with them (II/D.4-6). The distinctive feature of BDT is its reliance upon the structure of mathematical decision theory and probability theory to describe judgment/choice/decision/prediction/estimation behavior.

Of course, such an approach has obvious advantages in that the elegance and power of well established mathematical tools can be utilized in the description and analysis of observed psychological phenomena. The risk is that one's explanation of the psychological process under investigation will be distorted by an overzealous attempt to make the data fit the normative theory. Unfortunately, the reverse bias currently operates within other approaches (PDT, etc.) to give the impression that the axioms of mathematical decision theory and probability theory are rather useless as models of human judgment. The notion seems to be: "Let's see how bad a picture of cognitive functioning we can paint".

Perhaps a brief example is in order. Kahneman and Tversky's (1973) finding that people make non-regressive predictions which are "insensitive" to the reliability of the information upon which the information is based is a simple, yet "representative" arena within which to draw the contrast. Kahneman and Tversky (1973) hypothesize an internal cognitive mechanism to account for the observed result; subjects "predict by representativeness. That is, they select or order outcomes by the degree to which the outcomes represent the

essential feature of the evidence". In a later paper, Kahneman and Tversky (1977) describe, in detail, a debiasing procedures by which people can be cajoled, convinced and/or coerced into abandoning the representativeness heuristic in favor of a strategy which generates predictions which necessarily regress to the mean (See Table III-F-2).

From a BDT prespective, the finding is interesting, yet interpretable within the formal structure of mathematical decision theory. Regressive predictions (i.e. in which $\beta_x = r_{xy}$) are not an a priori requirement for normative prediction. The ordinary least squares (OLS) estimate of β is a consequence of the loss function chosen to formally represent the subjective loss or regret which results from imperfect prediction. The notion that an OLS estimator is somehow "God's" estimator is totally fallacious. In many real world situations, it is obvious that the loss function underlying OLS estimators are certainly not an accurate representation of subjective loss or regret.

One loss function, though not the only one to account for the "representative" predictions of Kahneman and Tversky, is that of summing the squares of perpendicular distances from each data point to the regression line (in contrast to the OLS procedure of summing the squares of vertical distances). Such a loss function always produces an estimate of β equal to plus or minus one, depending only upon the sign of the correlation coefficient, and not upon its magnitude. It seems reasonable in Kahneman and Tversky's experimental setting that S's would in fact attempt to match ordered pairs of x and y , and not be concerned with only matching y 's (as dictated by OLS). In effect,

Kahneman and Tversky have simply assessed each S's loss function. Their conclusion that S's are behaving in a grossly biased and sub-optimal manner is puzzling. Subjective loss functions (i.e. utility tradeoffs) are, after all, a property of the subject. In the absence of any explicit instructions in the form of costs or payoffs, any loss function is as good as an other. If one were to consider the class of all possible loss functions, it is clear that all regression lines with absolute slope between 0 and ~~+0~~ are possible. Regression lines with slopes ± 1 are like "average regression lines" over the space of all possible lines; Not a wholly unreasonable approach on the part of the subject.

REFERENCES

Kahneman, D. and Tversky, A. On the psychology of prediction.

Psychological Review, 1973, 80, 237-251.

Kahneman, D. and Tversky, A. Intuitive prediction: biases and corrective procedures. Research Report, Eugene, Oregon: Decision Research, 1977.

Intended uses of research results:

A number of "BDT-type" studies have investigated behavioral issues of utility elicitation. Early studies by von Winterfeldt and Edwards (1973) and Fisher (1972) started to answer questions of utility validation.

Since Gardiner and Edwards' application of SMART to public land use decisions in 1975, a number of studies have attempted to unravel difficult questions relating to number of dimensions employed, how weights should be ascertained, how utility measures should be determined, and a variety of other issues (See Fischer, 1976, 1977; von Winterfeldt 1977; Eustace, 1978)

In addition, problems of interpersonal conflict in probability assessment has been reviewed by Seaver (1976) and extensive empirical research has been completed on both behavioral and mathematical solutions to the problem (Seaver, 1978). The question of interpersonal conflict in utility assessment is the topic of a recent empirical study by Eils (1977) in which the decomposition strategy of MAUT was compared to a holistic choice approach under two different behavioral procedures for resolving interpersonal conflict. In short, much empirical research has been accomplished, most of it since Gardiner and Edwards 1975 paper, to lend support to the techniques and procedures of elicitation required by SMART.

REFERENCES

- Eils, L. C. III. The effects of communication strategy and decision technology on the process and product of decision making groups. University of Southern California (Unpublished dissertation), 1977.
- Fischer, G. W. Multidimensional utility models for risky and riskless choice. Organizational Behavior and Human Performance, 1976, 17, 127-146.
- Fischer, G. W. Convergent validation of decomposed multi-attribute utility assessment procedures for risky and riskless decisions. Organizational Behavior and Human Performance, 1977, 18, 295-315.
- Seaver, D. A. Assessment of group preferences and group uncertainty for decision making. Research Report 76-4, Social Science Research Institute, University of Southern California, 1976.
- Seaver, D. A. How groups can assess uncertainty: Human interaction versus mathematical models. Unpublished dissertation, The University of Michigan, 1978.
- von Winterfeldt, D. and Edwards, W. Evaluation of complex stimuli using multi-attribute utility procedures. Technical Report, The University of Michigan, Engineering Psychology Laboratory, No. 011313-2-7, October, 1973a.
- von Winterfeldt, D. Experimental tests of independence assumptions for risky multiattribute preferences. Research Report 76-8, Social Science Research Institute, University of Southern California, 1977.

Is IIT Simply a Special Case of SJT?

No one would question that IIT and SJT research stem from different origins and exhibit divergent theoretical positions. But what about methodology? McClelland points out that "The major difference is that IIT uses ANOVA instead of regression designs (p. III/E/4)." It may be useful to examine the relationship between SJT and IIT further, taking into account "that linear multiple regression analysis and the analysis of variance (and covariance) are identical systems (Cohen, 1968)."

Of course, the algebraic equivalence of the two is not the crucial issue. Mathematical statisticians have long known that both are a special case of the general linear model; Psychologists have become more and more cognizant of this fact in the ten years since Cohen's (1968) article in Psychological Bulletin. McClelland's distinction is in reference to the constraints the ANOVA paradigm places on IIT research. It is generally recognized that multiple regression (MR) is more general than ANOVA (or ANCOVA) (Cohen, 1968; Kerlinger and Pedhauzer, 1973). Every ANOVA (or ANCOVA) problem may be formulated as a MR problem, but the reverse is not true. Thus, from a methodological perspective, IIT is but a special case of SJT!

As an example, consider a recent paper by Zalinski and Anderson (submitted for publication). Subjects rated two, three, and four-factor job descriptions "by inserting a pin in a 200-mm graphic rating bar." The dimensions were managerial supervision, co-worker

cooperation, pay and opportunity for promotion, enjoyment and rewards of the work itself. Each dimension was evaluated at three nominal levels (LO, MED, and HI), resulting in 81 stimulus alternatives; Each subject made all 81 judgments three times. Averages over the three responses were entered into a standard 5-factor ANOVA design with one observation per cell (3(Management) x 3(Co-worker) x 3(Pay) x 3(Work enjoyment) x 20(Subjects)). Treating subjects as a random factor, all main effects and interactions of the first four factors were tested by comparing the mean squared error for the effect (MSE_E) to the mean squared error of the effect by subject interaction (MSE_{ExS}). For example, to test the Management (M) by Pay (P) interaction effect, the following F ratio was computed: $F = \frac{MSE_{PxM}}{MSE_{PxMxS}}$.

Early regression approaches to the "repeated measures" case of ANOVA were admittedly cumbersome, but new and innovative coding techniques by Pedhazur (1977) make the above problem a rather straight-forward application of MR. Since the original data would be necessary to demonstrate the required coding scheme, it is not presented here (see Pedhazur (1977) for three good examples of the necessary coding procedure). Testing the main effects and interaction effects reduces to simply testing the significances of differences between multiple R's for appropriate sets of coded independent variables.

An SJT researcher would probably carry the above analysis out with two slight variations: 1) MR's would be performed on an individual subject basis (using the 3 replications/subject as the random fifth factor), and 3) the magnitude of effects would probably be estimated, but not tested for statistical significance.

These two points are more a matter of taste (albeit an important one) which do not reflect any sort of limitations on the experimental paradigm or mathematical model used by SJT researchers. These two differences in procedure are of interest, but probably not critical. SJT researchers are aware of two simple facts: 1) when individual differences are prevalent (as they usually are), nomothetic analyses are of questionable value (see McClelland's section III/B

and Edwards's comments); and 2) Hypothesis testing almost always serves no useful purpose (Edwards, 1965; Edwards, Lindman and Savage, 1963).

IIT researchers are at least partially aware of these facts, since ANOVA results are often appended with qualifications asserting that response patterns for individual subjects demonstrated similar trends (as noted by McClelland, III/B/9) and that significant hypothesis tests may account for very little variance (Zalinski and Anderson, submitted for publication). It is probably safe to conclude that much of what IIT researchers do within the ANOVA paradigm is do-able within the MR approach of SJT. SJT offers the additional flexibility of studying correlated independent variables with continuous dimensions.

Of course, IIT is much more than just the ANOVA paradigm. It should be mentioned (since it isn't in the Colorado Report) that many of the other focal topics of IIT research are not well understood by those outside of IIT. If a successful integration of approaches is to occur, some resolution of these issues must be sought. For example, Hodges (1973) demonstrated that "more complex adding models cannot be distinguished from averaging models" and concluded that "the discussion of averaging versus adding models is of questionable value in advancing our understanding of information integration". In addition, Schönemann, Cafferty, and Rotton (1973) proved two rather general theorems in support of their statement that IIT methodology of estimating weights and scale values simultaneously is invalid; "The weights....whether obtained by a 'graphical method' or by some 'optimal statistical procedure', are empirically empty parameters so that, in fact, the additive functional measurement model reduces to a special case of additive conjoint measurement". Anderson (1973) responded to these charges, but the issues are far from cleared up. Surprisingly little attention is paid to them by IIT researchers. Norman (1976a, b) has developed a clever estimation procedure which seems to deal with the Schönemann et al (1973) problem, but Zalinski and

Anderson state that the method may only be applied to single subject analyses. In light of the previous discussion, this may not be a severe restriction.

In conclusion, that portion of the IIT methodology which I understand (the ANOVA paradigm) appears to be a special case of the more general MR approach of SJT. An interesting innovation by Barron, F.H. and Person, H.B. (submitted for publication), should serve as a catalyst to convince even the most disbelieving skeptic. Their idea is to utilize non-symmetric orthogonal experimental designs from ANOVA to specify the attribute levels for a reduced number of holistic judgments. These judgments are then analyzed via regression techniques to obtain estimates of attribute importance. Their proposed method of assessing importance weights turns out to be a special case of both the ANOVA and MR approaches! (practically speaking, the Barron and Person idea seems to be superior to standard approaches in either IIT or SJT).

I confess that, in light of the Hodges (1973) and Schönemann et al (1973) arguments, I simply do not understand much of the IIT methodology which is not subsumed by SJT. These more advanced methodological issues of IIT research are not addressed within the Colorado Report.

R E F E R E N C E S

- Anderson, N.H. Comments on the articles of Hodges and of Schonemann, Cafferty, and Rotton. Psychological Review, 1973, 80, 88-92.
- Barron, H.F. and Person, H.B. Assessment of multiplicative utility functions via holistic judgments (submitted for publication).
- Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968, 70, 426-443.
- Edwards, W. A tactical note on the relation between scientific and statistical hypotheses. Psychological Bulletin, 1965, 63, 400-402.
- Edwards, W., Lindman, H., and Savage, L.J. Bayesian statistical inference for psychological research. Psychological Review, 1963, 70, 193-242.
- Hodges, B.H. Adding and averaging models for information integration. Psychological Review, 1973, 80, 80-84.
- Kerlinger, F.N. and Pedhazur, E.J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart, and Winston, 1973.
- Norman, K.L. A solution for weights and scale values in functional measurement. Psychological Review, 1976a, 83, 80-84.
- Norman, K.L. Weight and value in an information integration model: Subjective rating of job applicants. Organizational Behavior and Human Performance, 1976b, 16, 193-204.
- Pedhazur, E.J. Coding subjects in repeated measures designs. Psychological Bulletin, 1977, 84, 298-305.
- Schonemann, P. H., Cafferty, T., and Rotton, J. A note on additive functional measurement. Psychological Review, 1973, 80, 85-87.
- Zalinski, J. and Anderson, N.H. Measurement of importance in multiattribute utility models (submitted for publication).

Statistical Measures of Importance in SJT

A key idea of SJT is the notion of an objective statistical ~~measure~~ ^{method} of importance. In laboratory studies (i.e. MCPL), statistical weights are used as feedback to the S, either in the form of Task Information, (TI), (concerning the environmental side of the lens model) or Cue Utilization (CU) Information, (concerning the subject(ive) side of the lens model). In real world policy capturing research (i.e. bootstrapping), derived statistical weights are used to describe the subjective judgment policy of the expert. In either case, statistical measures of importance derived from a multiple regression analysis of holistic choices are relied upon to provide an "understanding" of the judge's cognitive process.

One of the most useful features of SJT is its flexibility to cope with situations in which factorial designs and zero intercorrelations among attributes are not representative of the system under study. If a set of attributes have all zero intercorrelations, one statistical measure of importance is about the same as any other. Unfortunately, when cues or dimensions are correlated, different statistical measures of importance may be totally contradictory. This point has been demonstrated at least three times over the last 16 years (Ward, 1962; Darlington, 1968; Schmitt and Levine, 1977). Darlington's (1968) conclusion is particularly pointed: "It would be better to concede that the notion of 'independent contribution to variance' has no meaning when predictor variables are intercorrelated." Though simply stated, the problem has not been solved and few SJT researchers have seriously addressed the issue.

A somewhat related problem, also ignored by most SJT researchers, is the issue of estimating regression weights (β 's) under conditions of high multicollinearity. It is well known that ordinary least squares (OLS) estimates of the β 's are highly unstable when attribute (cue) intercorrelations are high. That is, the OLS estimates may be quite different from the population parameters of the β 's, resulting in a substantial shrinkage in multiple correlation upon cross-validation. Recent discussions of the multicollinearity problem may be found in the psychological literature (Price, 1977; Darlington, unpublished manuscript).

One solution is to utilize biased estimates of β , via any one of several recently developed techniques. One of the more popular approaches is that of ridge regression, invented by Hoerl in 1962. In addition to the articles cited above by Price and Darlington, excellent discussions of ridge regression may be found in Hoerl and Kennard (1970a, b) and Marquardt and Snee (1975). Numerous simulation studies have generated evidence indicating that ridge estimates are superior to those of the more common OLS approach (Dempster, Schatzoff, and Wermuth, 1977).

Thus, even if a statistical measure of importance could be justified on theoretical grounds, problems would still arise if it involved estimates of the β 's e.g. β^2 or Hoffman's (1960) relative weights, $B \cdot r/R^2$. Statistical measures of importance involving unbiased estimates of regression weights may be irreparably misleading, at least for some of the purposes to which SJT researchers put them. Perhaps ridge estimates of β weights should be used when calculating statistical indices of importance.

Although the discussion has been restricted to SJT research, these issues bear heavily on current research effort in BDT to validate the subjective weighting procedures of Edward's "SMART" and the various procedures suggested by Keeney and Raiffa (1976) for eliciting weights for riskless additive multiattribute utility functions. In addition to the numerous convergent validation studies of BDT research cited elsewhere in these comments, Cook and Stewart (1975) provide an excellent demonstration of the convergent validity of subjective weighting approaches. As Schmitt and Levine (1976) pointed out, subjective weighting procedures have received a fair amount of bad PR. Early studies reviewed by Slovic and Lichtenstein (1971) tended to rely on correlations between sets of statistical weights and subjective weights to measure the degree to which the latter approximated the former. Since composites resulting from using totally uncorrelated sets of weights may be highly correlated, these studies are highly suspect. Recent research within the MCPL paradigm by Schmitt (1978) and John (unpublished manuscript), suggest that subjective weights are as valid in describing subjective policy judgments as statistical weights.

One problem in research attempting to validate subjective weights is that of determining which statistical indices of importance should be used in making the comparison. Ultimately, one isn't quite sure which procedure is validating which. Perhaps subjective weights should be used as a means of determining which of the statistical measures of importance is psychologically valid. As Schmitt and Levine

(1976) point out, "all the (statistical) indices are paramorphic in nature, i.e., they were derived from regression analyses which have nothing to do necessarily with the actual decision process."

Many unresolved issues relating to statistical weights have been presented under the assumption that they are critical to SJT and thus of some importance to the integration process attempted in the Colorado report. Unfortunately, the crux of the problem seems to be with the case of non-zero attribute (cue) intercorrelations; which is the very heart and soul of research within the lens model. As Edwards points out in his comments, there is perhaps little justification for modeling one environmental system over any other. Unless SJT researchers can address the problems inherent in statistical measures of importance, they may be stuck with Edwards' assertions, whether they agree with them or not.

REFERENCES

- Cook, R. L. and Stewart, T. R. A comparison of seven methods for obtaining subjective descriptions of judgmental policy. Organizational Behavior and Human Performance, 1975, 13, 31-45.
- Darlington, R. B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.
- Darlington, R. B. Reduced-variance regression (unpublished manuscript).
- Dempster, A. P., Schatzoff, M. and Wermuth, N. A simulation study of alternatives to ordinary least squares. Journal of the American Statistical Association, 1977, 72, 77-91.
- Hoerl, A. E. and Kennard, R. W. Ridge regression: applications to nonorthogonal problems. Technometrics, 1970, 12, 69-82.
- Hoerl, A. E. and Kennard, R. W. Ridge regression: biases estimations for nonorthogonal problems. Technometrics, 1970, 12, 55-67.
- Hoffman, P. J. The paramorphic representation of clinical judgment. Psychological Bulletin, 1960, 57, 116-131.
- Keeney, R. and Raiffa, H. Decisions with Multiple Objectives. New York: Wiley, 1976.
- Marquardt, D. W. and Snee, R. D. Ridge regression in practice. The American Statistician, 1975, 29, 3-19.
- Price, B. Ridge regression: applicaiton to nonexperimental data. Psychological Bulletin, 1977, 84, 759-766.
- Schmitt, N. Comparison of subjective and objective weighting strategies in changing task situations. Organizational Behavior and Human Performance, 1978, 21, 171-188.

Schmitt, N. and Levine, R. L. Statistical and subjective weights:
some problems and proposals. Organizational Behavior and Human
Performance, 1977, 20, 15-30.

Ward, J. H. Jr. Comments on "the paramorphic representation of
clinical judgment." Psychological Bulletin, 1962, 52, 74-76.

COMMENT BY
AMOS TVERSKY
STANFORD UNIVERSITY

Amos Tversky

My discussion of the Colorado report will be divided into two parts. First, I wish to comment on general conceptual issues involved in an attempt to integrate the area of judgment and decision, and second, I wish to react to some specific sections of the report.

Since the stated objective of the report is the integration of judgment and decision theory it is instructive to examine first the meaning of 'integration' and to what extent it is an achievable goal. Integration is used in science in several distinct senses. A strong sense of the term could mean reducing all theories in a given field to a common conceptual language that will enable one to treat them as a unified theory. Although such integration might be highly desirable, I seriously doubt whether it is achievable at this stage. Another sense of integration might refer to an attempt to facilitate the interaction between research programs, offer common interpretation to their findings, and increase the communality between different approaches. The latter goal strikes me as both feasible and worthwhile, and the Colorado Report represents a serious attempt in this direction.

There have been several efforts in the past to relate different approaches to judgment and decision making. Ward Edward's review paper from 1954 is perhaps the first successful attempt of this sort in which Edwards underscored the mutual relevance of economic theory and experimental studies of decision making. The Slovic and Lichtenstein review from 1971 represents another attempt of this sort in which the authors not only reviewed the experimental and theoretical work in the field but also attempted to reconcile and discuss the differences between the Bayesian and regression approaches to judgment and decision research. A third effort is represented by the forthcoming book by Richard Nisbett and Lee Ross which attempts to integrate much recent research on judgment with

relevant work in social psychology. Although these manuscripts differ in their focus and orientation, they all share the objective of providing an integrative view of the relevant sub-fields. The Colorado Report differs from the above in that it focuses more on conceptual issues and research strategies, and less on review of experimental findings. Following this orientation let me discuss first some of the key features of contemporary research on judgment and decision making that contribute to its theoretical promise and its practical relevance.

The presence of normative theory

Much of the interest in the psychological analysis of decision and judgment, in my opinion, stem from the presence of a well-articulated and generally accepted normative theory to which behavior can be compared. Many researchers, notably Egon Brunswick and Herbert Simon, emphasized the importance of an adequate analysis of the nature of the task and the environment in which behavior takes place. Thus, the field of vision, for example, owes a great deal to optics, and psycholinguistics benefits greatly from the linguistic models of grammar. In the same way, the presence of a normative theory of belief (i.e., Bayesian statistics) and of action (i.e., expected utility theory) provide powerful tools for the analysis of judgment and decision.

The contrast between normative and observed behavior

One could easily imagine a world in which behavior is adequately described by the Bayesian model, and in which the descriptive and the normative theories coincide. In such a world, there would be no need for a psychological theory, apart from the normative theory of judgment and choice. Our world is certainly not like that. Human behavior departs from the rational model in many ways and hence the study of judgment and decision is characterized by the presence of tension between the predicted rational behavior and the observed human behavior. The departures from the normative model could result from limited information capabilities, originally suggested by Simon and later documented and expanded by Slovic and Lichtenstein, or it could represent basic disagreements between

the calculus of chance and the psychological conception of uncertainty as suggested by Daniel Kahneman and myself. Be that as it may, the attempt to understand, describe and perhaps reduce the discrepancy between prescriptive and descriptive theories of choice is characteristic of much theoretical work in the field.

Applications

One key feature associated with contemporary work on judgment and decision, which sets it apart from many other areas of cognitive psychology, is the concern with applications. The fact that many important decisions and inferences in the real world are carried out in an intuitive manner suggests that the biases and fallacies observed under controlled conditions may operate in the real world as well. This hypothesis is supported by much research about clinical diagnosis, selection procedures, and market behavior. The relevance of research on judgment and decision to the world of reality, and the potential for improving the quality of judgments and decisions have made the field much more exciting.

The concern with applications is shared by most participants of the conference, as made evident by the Colorado Report. There is no general agreement, however, regarding the manner in which psychological knowledge could be utilized to improve inference and decision. Nevertheless, two general approaches appear to emerge: a technological approach and an educational approach. The technological approach to decision analysis, represented by the work of Ward Edwards and others, is concerned primarily with the development of feasible procedures designed to improve decision making. On the other hand, the educational approach, advocated by Daniel Kahneman and myself, attempts to make decision makers aware of their own errors and biases, so that they can be reduced or eliminated. The two approaches are clearly complementary rather than competing, and much applied research is required before they could be adequately assessed.

Having discussed some of the general characteristics of the field, let me turn now to list what I perceive to be the major outstanding problem in the field

from both normative and descriptive standpoints.

The decomposition of multi-attribute structures.

Although this subject has been the topic of much theoretical analysis (e.g., the book by Keeney and Raiffa) and much empirical research, we do not know yet how people combine values or determine exchange rates between different attributes, and what mathematical structures would best represent their behavior. Furthermore, although recent theoretical work has revealed the mathematical structure underlying multi-attribute models, the normative question regarding the adequacy of any particular multi-attribute structure has not been resolved in a satisfactory manner. Unlike the case of decision making under risk where probability theory provides the key to the normative analysis, there is no comparable normative structure for the combination and updating of utilities.

The combination of evidence

Much empirical and theoretical work has addressed the issue of how people combine evidence from different sources and revise their opinion on the basis of new data. Nevertheless, many central problems have not yet been solved. One outstanding problem is the characterization of the relation of conditional independence between items of evidence, so as to determine how they should be combined. Shafer's work on the mathematical theory of evidence represents one interesting new direction of normative research. A second open problem concerns the resolution of incoherence. Suppose one discovers that his probabilities are mutually inconsistent. How should they be revised? The latter problem exemplifies the type of research emerging from decision analysis that has both descriptive and prescriptive elements.

Basis for application

Most students of decision making believe that human judgment could be improved by analytical thinking or formal modelling. The basis of this belief, however, is yet to be established. Although most decision analysts feel that their work helps their clients in a variety of ways, systematic evidence on this

point is lacking. Since direct experimental tests of decision analysis are difficult to obtain, the development of a program of research that would shed light on the usefulness of decision analysis represents a major challenge for students of the field.

Having commented on some general issues raised in the Colorado report, let me turn now to more specific comments regarding the discussion of Psychological Decision Theory with which I am identified. In general, the discussion of PDT in the report strikes me as highly cogent. In my response, I wish to discuss first the status of theoretical concepts such as representativeness, and then comment on some methodological issues raised in the report.

On the Status of Concepts

The concepts of representativeness and availability have been criticized (in the report and elsewhere) for lack of rigorous definitions. I do not wish to discuss here the general problem regarding the usefulness of intuitive concepts in scientific inquiry. I only wish to clarify the status of representativeness and availability and to explicate the logic that underlies their use. To begin with, we assume that entities or events usually vary in the degree to which they are representative of some prototype, model, or data generating process. This assumption is supported by the observation that people order entities (e.g., sequences of heads and tails, or thumbnail descriptions of individuals) according to representativeness in a reliable and consistent manner. Similarly, we assume that occurrences or instances vary in their availability: some instances come readily to mind while others are difficult to recall. Thus, the representativeness of an object to a given category and the availability of an instance in a particular context are viewed as attributes that can be directly judged or assessed.

The representativeness hypothesis asserts that judgments of the likelihood of events are biased by their representativeness. Stated differently, we hypothesize that representativeness and availability are used as cues or heuristics in judgments of uncertainty. It is important to distinguish between representativeness or availability that refer to perceived properties of stimuli, and the representativeness or availability heuristics that relate these characteristics of the stimuli to judgments about their likelihood. The representativeness heuristic, for example, is supported by the observation that subjects who judge a description of Mr. X to be more representative of a librarian than that of a farmer also believe that Mr. X is more likely to be a librarian rather than a farmer -- with little or no regard to the base rates of these classes.

The representativeness and the availability hypotheses, therefore, concern the relation between two conceptually distinct responses (e.g., similarity ratings and likelihood judgments), and no additional assumptions are needed in order to test them. In some studies, however, we did not actually collect representativeness or availability information because we felt (erroneously perhaps) that the relations are self-evident. For example, we assumed that people regard the sequence HTTHTH as more representative than the sequence HHHTTT. This assumption, however, can be readily put to an empirical test - if its validity is questioned. Hence, the reliance on the readers' intuitions and the use of plausibility arguments were introduced to simplify the experiments, and they can usually be replaced by judgmental data.

The role of representativeness in judgments of uncertainty, therefore, can be investigated experimentally - even in the absence of a comprehensive theory that explains its determinants. (The same argument applies to availability and causality.) The question as to what makes an object

representative of a given category is a central problem in both philosophy and psychology. Recent research in Cognitive Psychology has demonstrated the importance of prototypicality or representativeness in perceptual learning, semantic memory and the formation of categories. The development of a full-blown theory to explain the determinants and the assessment of representativeness is a challenging task, and I have recently proposed an approach to the problem that is based on the notion of a feature-matching function. Adequate theories of representativeness or causality would undoubtedly shed light on many aspects of human judgment. However, such theories are not required in order to demonstrate that representativeness or causality bias judgments in a predictable manner.

Methodological Considerations

The general orientation of PDT is ideographic rather than nomothetic. Between-subjects designs and aggregate data are used to achieve experimental control and to simplify the demonstration. Consider, for example, the certainty effect which underlies prospect theory. This hypothesis asserts that a person who is indifferent, say, between a cash gift of \$X and a 50-50 chance to win \$1000 or nothing, will prefer 5% chance to win \$1000 over 10% chance to win \$X. For a direct test of the certainty effect, therefore, we must first find the appropriate value of X for each individual, and then test the predicted preference. Alternatively, one could select some value of X, say \$450, that is preferred to $(1000, \frac{1}{2})$ by most subjects, and then test whether $(1000, \frac{1}{20})$ is preferred to $(450, \frac{1}{10})$. Because of individual differences, the predicted pattern is not expected to hold for all respondents -- even when the certainty effect is satisfied without exceptions. Hence, the certainty effect is supported by the observation that most subjects selected the predicted option in both choices,

as was indeed the case.

Much of the data associated with PDT (particularly the work conducted by Kahneman and myself) were obtained from questionnaires presented to moderate-sized groups, where each subject answered only a small number of questions. This methodology differs drastically from, say, the typical studies of Bayesian updating in which each subject made many responses under similar conditions. Our choice of research strategy was guided by the fear that the responses to many repetitive questions do not add significant new knowledge, and by the hope that the answers to a few, simple, carefully worded questions will provide useful information about values and beliefs. Naturally, the choice of method is determined both by the objectives of the study and by many practical considerations.

COMMENT BY

NORMAN ANDERSON

UNIVERSITY OF CALIFORNIA, LA JOLLA

Norman Anderson

Information integration theory has four main characteristics:

0. Study of Multiple Causation
1. Cognitive Algebra.
2. Functional Measurement.
3. Empirical Foundation.

Multiple causation might be called the basic problem of psychology; certainly it is for all the approaches represented at this conference. But multiple causation is hard to study. When diverse causal forces are at work, each pushing in its own direction, the observed net resultant is difficult to predict and difficult to analyze. I believe it is fair to say that integration theory provides the first effective, general attack on the problem of multiple causation--at the psychological level of cognitive process.

Cognitive algebra forms the foundation of this approach. Numerous investigators in almost every area of psychology have conjectured that human judgment may obey simple algebraic models. But without a capability for psychological measurement, these conjectures have remained hopeful verbalisms. Functional measurement has made it possible to test these conjectures.

Happily, many of these algebraic models have been shown to be correct, a unique contribution of information integration theory. Cognitive algebra has been found to operate in many areas of decision theory, including subjective probability and intuitive statistics, as well as in person perception, attitudes and attributions, psychophysics, and psycholinguistics. This extensive empirical work provides the substantive

foundation that is necessary to progress from paper promises to proper theory.

Beyond its intrinsic interest, cognitive algebra has various practical applications. As one example, it can provide a validational base for self-estimated parameters for use in multiattribute models.

ILLUSTRATIVE APPLICATIONS

1. Probability Models: Irrational Betting

Lola Lopes' (1976) thesis on poker playing provides an important illustration of the methods and results of information integration theory. Experienced poker players played a modified game of 5-card stud against two computerized opponents. On each hand, they made an even-money side-bet with the computer that they could beat both opponents. After the subject bet, the computer turned over the hole cards and collected from or paid off to the subject.

Lopes wished to test the multiplication rule:

Subjective Probability (Beating both A and B)

= Subjective Probability (Beating A)

x Subjective Probability (Beating B).

This rule assumes independence, an assumption that was facilitated by the instructions. The functional measurement test of this multiplication rule is simple: If it holds, then the data will plot as a linear fan. The subjects' judgments of probabilities did indeed follow this linear fan pattern; so did their betting behavior.

The linear fan pattern is clear in Figure 1. The vertical axis plots amount bet (between 1 and 30 cents) as a function of hand strength of the conservative opponent (curve parameter) and hand strength of the average opponent (horizontal axis). The linear fan pattern implies that betting behavior followed a simple cognitive algebra.

This betting behavior is quite irrational, of course. When subjective probability is less than 1/2, the minimum should be bet; when subjective probability is greater than 1/2, the maximum should be bet. Rationally, the curves should be step functions, not straight lines. The suboptimal betting strategy can cost subjects as much as 12-1/2%.

2. Other Probability Models

Functional measurement has opened up the road for an attack on other probability models. Classic work by Shanteau has provided an empirically validated foundation for SEU theory. This work is too well-known to need mention but it is of interest that the key to success lay in a multiplying rule. The additive rule, which had dominated previous work, has suffered repeated failures.

An interesting application by Wyer found good support for what may be called the probability averaging model,

$$\text{Prob}(A) = \text{Prob}(B) \text{Prob}(A|B) + [1-\text{Prob}(B)]\text{Prob}(A|\bar{B}),$$

as shown in Figure 2. It should be added that Wyer reported that certain other probability models did not work too well.

3. Weak Inference

Multiplying rules are frequent in decision and judgment, as with the above probability models. For practical prediction, it may be useful

to replace them by adding models or by multiple regression models. Some caution is necessary, however, as Figure 3 from Anderson and Shanteau (1977) demonstrates. The "observed" points represent areas of rectangles. They form a linear fan as required by the multiplying model, $\text{Area} = \text{Base} \times \text{Height}$. The four heavy parallel lines represent the best predictions of the linear model. The correlation between observed and predicted is .885; that might seem to reflect good performance by the linear model. However, a glance at the large discrepancies in the graph seems rather to support Tukey's position that the correlation coefficient can be "an enemy of good data analysis."

This point is hardly novel, to be sure, but it has been ill appreciated. Much work in judgment and decision has little value because it rests on such "weak inference" methodology.

An experimental example from consumer attitudes in Figure 4 reemphasizes this point. In the top panel, the parallelism of the solid curves and the crossover of the dashed curve imply that product attributes are integrated by a configural (averaging) rule. But suppose that this fact is ignored and a linear model is fit to the data. Clearly, this fit has some serious errors because it requires five parallel lines whereas the dashed line is markedly nonparallel. But little hint of this appears in the scatter plot of predicted vs. observed in the lower panel, or in the correlation of .980.

The great practical usefulness of the linear model goes without saying. But for the study of cognitive processes it may rather be harmful.

4. Self-Estimated Parameters

Because it has solved the problem of psychological measurement, integration theory provides a basis for developing techniques for self-estimated parameters. Many if not most practical applications of multi-attribute models, for example, performance depend on such direct assessment. To say that current assessment methods are arbitrary does not mean that they are not reasonable or useful; it does mean that they need to be evaluated and improved.

The concept of weight, noted below, is one example. Our work on job satisfaction suggests that people are fair judges of their own weights and we hope to develop improved techniques for this purpose.

A second illustration is provided by the controversy between the common method of ratings and S. S. Stevens' method of magnitude estimation. Stevens has tirelessly reiterated his message that ratings are the worst form of direct measurement and that magnitude estimation provides the true measure of psychological value, not merely in psychophysics but in social judgment as well.

Functional measurement provides a validation criterion to resolve this controversy. Ratings and magnitude estimation have equal opportunity. An extensive network of evidence has favored the rating method whereas magnitude estimation has generally failed. A vital aspect of this work that deserves present emphasis was the development of procedures that largely eliminate the well-known biases that can afflict the rating method.

5. The Concept of Weight

The Colorado Report (II/E/8) suggests that the concept of weight is fundamental in all six approaches. In most applications, however, the

weight measure is arbitrary and sometimes meaningless. Integration theory provides a theoretical foundation for the concept of weight as amount of information.

This issue is important for multiattribute theory in which people give self-estimates of importance or weight. Are these self-estimates valid? Most studies of this issue conclude that people are very poor at estimating the weights that are functional in their judgments (see reviews by Slovic & Lichtenstein, 1971; Slovic, Fischhoff, & Lichtenstein, 1977). But this problem cannot be answered without a validity criterion, and perhaps without exception the choice of validity criterion has been arbitrary.

Multiple regression approaches illustrate this arbitrariness quite well. As one problem, the regression weights are confounded with the unit of their scale of measurement and hence not generally comparable across variables. (To illustrate this point, the question of whether temperature is more important than humidity in industrial comfort doesn't make much sense when opposite answers can be obtained depending on whether temperature is measured in Celcius or Fahrenheit degrees.) Standardization does not solve the problem. A second, well-known problem is that weight estimates can be systematically erroneous when the predictor variables are intercorrelated.

Information integration theory provides a solution to this problem. The concept of weight, as distinct from scale value, is made clear in the averaging model, and the empirical success of this model provides a substantive foundation for a psychological concept of weight. Most important

for present purposes, this provides a basis for developing valid direct measurement of weights.

6. Unique Situations: Analysis with Self-estimated Parameters

Many situations are unique in the sense that they must be analyzed as they stand, without any kind of experimental or even environmental variation. Many applications of multiattribute utility theory are of this kind. Two applications from information integration theory are shown in the graphs below.

Figure 5 shows data from an experiment on attitudes toward U. S. presidents developed in group discussion (Anderson & Graesser, 1976). The cues included the opinions of the other members of the group, and each subject estimated the weight and scale value of each such cue in his/her own final attitude. These self-estimated parameters allow a prediction from averaging theory for a single discussion. The graph shows the fit between predicted and observed; the discrepancies were not significant.

Figure 6 shows date-preferences by females of males described by a photograph (Shanteau & Nagy, 1975). It was assumed that the photograph would elicit two cognitive responses: one of physical attractiveness, the other of the probability that the male would go out with them. Since these cognitive responses, or cues, are not under experimental control, self-estimated parameters were obtained from each subject. The theoretical predictions, obtained from a multiplying model, seem quite good.

These two studies were possible only because of prior developments in cognitive algebra and with numerical response methods. They illustrate the potential for a theoretical foundation for multiattribute models.

THEORETICAL COMPARISONS

Attribution Theory

As I have detailed elsewhere (Anderson, 1974, 1978), information integration theory provides a unified theoretical approach to attribution, both social and nonsocial, causal and noncausal. The need for a unifying theory was recognized by Jones, Kanouse, Kelley, Nisbett, Valins, and Weiner (1972, p. x) who say, "At present attribution theory is an amorphous collection of observations about naive causal inference," an appraisal reiterated by Jones and McGillis (1976, p. 390). In the present view, attributional judgments are not essentially different from the other kinds of judgments that have also been studied in integration theory.

Jones and McGillis discuss some of the limitations of and relations between Jones' correspondent inference theory and Kelley's analysis of variance cube. I feel that integration theory has certain advantages in generality, conceptual clarity, and analytical precision. This point may be briefly illustrated by reference to the two basic operations of integration and valuation.

Causal schemes. The concept of causal schemes provides a good basis for comparison. In the integration-theoretical approach, causal schemes are integration rules, operative cognitive structures that process input information to provide the desired judgment.

Suppose that two causes, F and G, give rise to some action or event, A. Judgments about the action represent forward inference, and judgments about the causes represent inverse inference. Symbolically,

$$A = F \otimes G,$$

$$G = A \otimes' F,$$

where \otimes and \otimes' are generalized integration operations. Kelley's response patterns of discounting and augmentation can be represented as special cases of these integration operations.

As an experimental illustration, it has been shown that

$$\text{Performance} = \text{Motivation} \times \text{Ability}$$

$$\text{Motivation} = \text{Performance} - \text{Ability}.$$

These attributions thus exhibit a simple cognitive algebra although this is not a mirror of mathematical algebra. Similarly, the well-known McArthur experiment can be nicely unified around the integration rule,

$$\text{Fear} = \text{Motivation} \times \text{Valence}$$

$$= \text{Person} \times \text{Environment}.$$

This view of causal schemes as integration rules is simple and natural and it has great analytical power.

Kelley's view is quite different. In his 1972 chapter, Kelley's definition reads thus: "A causal scheme is an assumed pattern of data in a complete analysis of variance framework." It might seem that Kelley had in mind a factorial design data table. That would make a causal scheme an actuarial table, conceptually quite different from a cognitive integration operation.

However, in his experiment, it becomes apparent that Kelley defines causal scheme as a specific array of stimulus cues. Kelley assumes three basic dimensions of information (distinctiveness, consensus, and consistency), and specific patterns of high and low values on these three

basic dimensions constitute the "causal schemata": "As such, the three patterns enter into the inferential process itself. Information is compared with them and is interpreted in terms of the pattern(s) with which it is consistent" (Orvis, Cunningham, & Kelley, 1975, p. 606). The HHH pattern, for example, would lead to an attribution to the environment as cause.

In the present view, the basic dimensions are the causal forces themselves, Person and Environment in the above example. Kelley's consistency dimension is not basic because such information will affect both causal forces. Integration theory also allows for the many kinds of pertinent information that do not lie on Kelley's three dimensions. Beyond that, it provides a basis for construct validity for the theoretical conceptualization.

Bayesian approach. Workers in social attribution have been remarkably insulated from general judgment theory (e.g., Fischhoff, 1976; Colorado Report, pp. I/3, II/E/27). Only recently have they begun to take cognizance of the Bayesian movement which has centered on the focal attribution problem of inverse inference. Unfortunately, the major attempt (Fishbein & Ajzen, 1975) to relate these two areas has thrown out the baby but kept the bathwater. This report suffers severely, as Fischhoff and Lichtenstein (in press) have pointed out, from elementary misconceptions about Bayesian theory. And it also fails to recognize that Bayesian theory has long been known to be invalid as a cognitive theory of psychological process. To this may be added that the Bayesian model is limited to probabilities and cannot handle general social judgment.

The Bayesian approach does have two appropriate uses. One is as a normative baseline. The other, more important I believe, resides in the conceptual framework of ideas such as diagnosticity, inverse inference, etc., quite apart from any particular model. In the recent rush to emphasize the numerous ways that people deviate from the Bayesian model, the truly pioneering contribution of Ward Edwards in bringing this conceptual framework to general attention has often been neglected.

Valuation. Where integration theory is weak and attribution theory is strong is in the study of the valuation operation. This refers to the processes that control the parameter or function values in the integration model. Thus, basic experimental work by Jones and his associates has been concerned with informativeness or informational content. Current concern over usage of base rate information, consensus and distinctiveness information, etc., can also be viewed in terms of valuation processes.

The integration models as they stand can make two valuable contributions to the study of valuation. One is that valuation operations often turn out to be prior integration operations (the "first integration problem" in attribution theory, Anderson, 1978). The other is that functional measurement can determine the functional values that are operative in any given situation. These functional values provide boundary conditions on the valuation operation. Any theory of valuation must obey these boundary conditions which, therefore, constitute a validation criterion that might not be otherwise attainable.

Lens Model--Multiple Regression Approach

The Colorado Report suggests that a unification of integration theory and the lens model approach may be facilitated by their quantitative

form, but perhaps precluded by different metatheories. I certainly agree that this matter should be explored but I am concerned because the view apparently being attributed to integration theory, on page II/E/28 and elsewhere, is one that I categorically reject. In particular, cue interaction is well recognized in integration theory, as in the concepts of inconsistency and redundancy. Also, we have long been concerned with configural integration.

The source of this misunderstanding, I suspect, lies partly in a confusion between the cognitive integration rule (which operates on a single stimulus presentation) and the frequent use of a factorial design (which is merely a diagnostic tool to discover the structure of cognition). In any case, the two following issues seem to be fundamental in any attempt at comparison.

Measurement. The lens model ordinarily requires scaled, numerical cues. In some cases, these are appropriately measured on physical scales, for example, objective probability and dollar value. No problem arises when the physical metric is appropriate to the task.

But how can this approach handle cues for which there is no physical metric? Subjective probability and utility, for example, as in Lopes' poker experiment. Or physical attractiveness as in the dating studies of Shanteau and Nagy. How can the lens model ascertain the functional, psychological values?

In these cases, measurement in the lens model appears to be arbitrary. That does not mean it is unreasonable or even necessarily incorrect. Arbitrary, commonsense measurement can often be useful. However, it can hardly serve as a basis for cognitive theory.

Model. No less arbitrary is the lens model itself. It apparently is a basic postulate, not to be questioned. At least, that seems to be the meaning of the statement on page III/E/6, bottom (and similarly on page III/E/8), that tests of goodness of fit are not used in the lens model approach "because it claims to be only a descriptive theory and the fact that all the judgments came from one individual is sufficient justification for describing those judgments in terms of weights and function forms."

How would this apply to Lopes' poker experiment? As I understand it, the usual regression analysis would be employed to fit an additive-type model to the data. The R^2 would obviously be high and so it might seem that the lens model gave a satisfactory description of the data. But the psychological process is actually multiplicative as shown so clearly by the linear fan pattern above. It would seem, therefore, that the lens model does not represent the cognitive process.

That is not an objection unless the lens model regression claims to be a cognitive theory. On page III/F/9, however, we are told that this approach is a cognitive theory. This apparent inconsistency needs clarification.

Multiplying rules arise frequently in decision theory, both normatively and descriptively. Integration theory has found evidence for these and other rules of cognitive algebra in numerous situations. These provide a base of evidence against which any attempt at cognitive theory must be assessed.

Task, Environment, and Optimality

The Colorado Report has numerous comments to the effect that integration theory neglects the structure of the environment and is indifferent to problems of optimality. It is true that we have been primarily concerned with cognitive structure. I was prepared to admit the criticism, therefore, and pledge to work toward broader applications. Certainly there is nothing in integration theory that precludes the study of optimality or task-environment structure. Indeed, our formulation can be applied to study environment structure in the same way as it has been applied to cognitive structure.

On reflection, however, it seemed that we had actually been doing quite a bit of this. Here are a few examples.

(a) In Lopes' poker experiment above, it was a matter of keen interest that the behavior followed a simple cognitive rule but not the rational rule. Moreover, this cognitive rule was a subjective analog of a normative probability model.

(b) In Wyer's study, similarly, the hypothesized model comes from normative theory and functional measurement allows an assessment of whether the algebraic structure is valid while allowing for subjective probability.

(c) Work by Shanteau and his associates on livestock judging deals with optimality in two ways: with respect to the factual criterion for the judgments; and with respect to capacity to integrate a multiple-cue stimulus field.

(d) In now classic work, Shanteau has also studied the "subadditivity" effect in SEU theory, also a deviation from normative theory.

(e) In related work on Bayesian theory, Shanteau found evidence for inadequate use of diagnosticity information, a theme also taken up by Kahneman and Tversky.

(f) In still more work comparing information integration theory and Bayesian theory, Shanteau found that obviously nondiagnostic information affected causal inference.

(g) In judgments of value of information for decisions, Shanteau and Anderson found an interesting suboptimality in which subjects followed a simple multiplying rule even though it cost them money.

(h) An example from legal psychology is the finding of Kaplan and others that legally irrelevant factors such as physical appearance can bias judgments of judges and juries.

(i) Finally, an extremely interesting use is adumbrated in the seminal work of Ebbesen and Konečni on bail setting. Here the judge's decision in the calm of his chambers provided, at least in my own view, his standard of justice against which to assess his decisions in the actual courtroom situation.

We should do more of this and hope to improve, but we have been primarily occupied with developing a theory of cognitive organization rather than of task-environment organization. At the same time, I suggest that those who have been more immediately concerned with suboptimality have been missing valuable opportunities in neglecting to use our approach for discovering covert processes underlying overt judgments. As Lopes' report demonstrates, and as several of the other articles indicate, deviation from optimality often reflects the operation of a simple cognitive algebra. Knowledge of the underlying process may provide the most effective path to understanding and correcting suboptimality.

References

In order to get this discussion draft to the conference in time, references have been omitted. Most of the cited references can be found in the Colorado Report or in the following four chapters.

Information Integration Theory: A Brief Survey. In D. H. Krantz, et al (Eds.), Contemporary developments in mathematical psychology (Vol. 2). Freeman, 1974.

Cognitive Algebra: Integration Theory Applied to Social Attribution. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 7). Academic Press, 1974.

Progress in Cognitive Algebra. In L. Berkowitz (Ed.), Cognitive theories in social psychology. Academic Press, 1978.

Integration Theory Applied to Cognitive Responses and Attitudes. In R. E. Petty, T. M. Ostrom, & T. C. Brook (Eds.), Cognitive responses to persuasion. McGraw-Hill, in press.

Figures

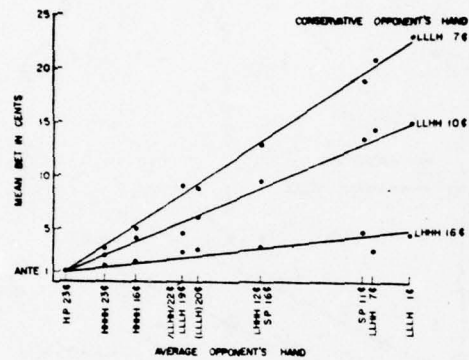


Figure 1. Obtained and predicted data for Part 1 of Experiment 3: Rest of the multiplicative integration rule.

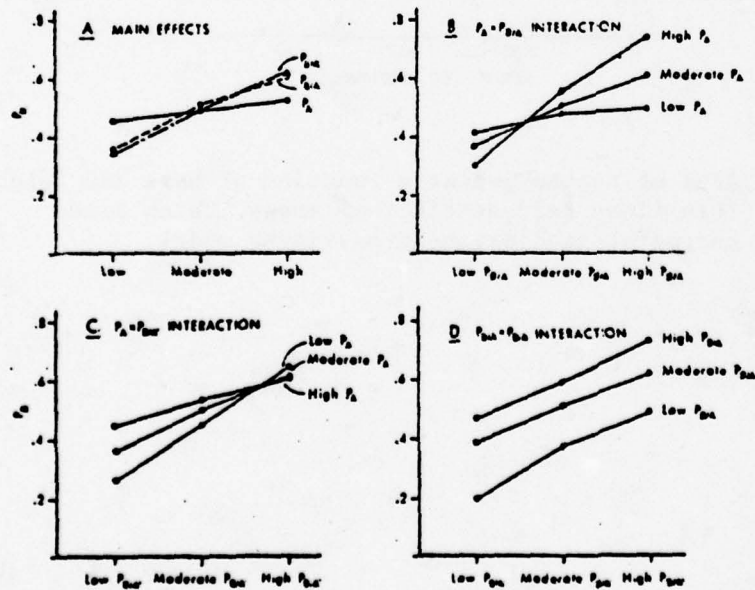


Figure 2. Main effects and interactions pertaining to P_B as a function of P_A , $P_{B/A}$, and $P_{B/A}$. (Wyer, 1975)

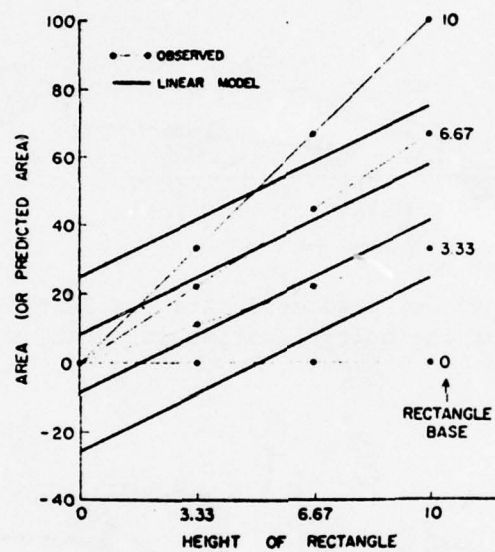


Figure 3. Area of rectangles as a function of base and height. Thin lines represent actual areas, thick lines represent predictions from linear model.

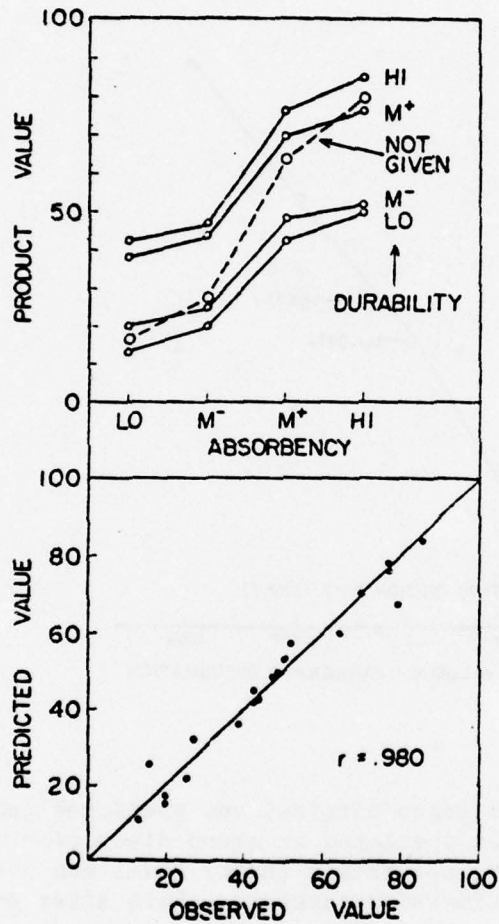


Figure 4. Upper panel gives factorial plot of mean judgment of product value as a function of absorbency and durability. L, M⁻, M⁺, and H denote stimulus values graded from low to high, respectively. Crossover interaction supports the averaging hypothesis, rules out an additive model. Lower panel plots predictions from additive model as a function of the observed means in the upper panel. (Data reanalyzed from Troutman & Shanteau, 1976.)

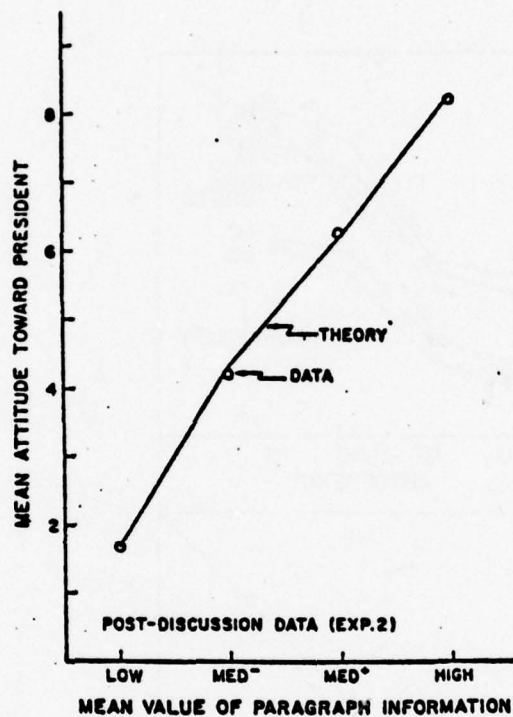


Figure 5. Observed (open circles) and predicted (solid lines) attitudes developed in group discussion. Predictions based on integration theory model and use of self-estimated parameters. Data after Anderson and Graesser (1976).

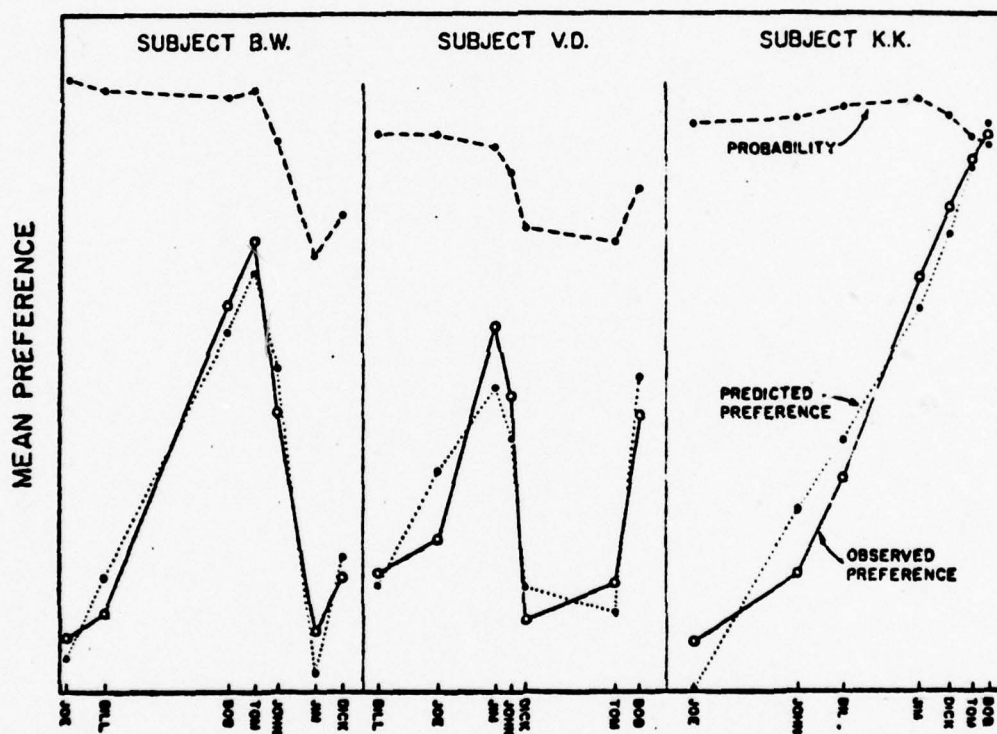


Figure 6. Data preferences of girls for boys as a function of physical appearance, each boy listed on the horizontal axis having been described by a photograph. Solid and dotted curves compare observed and predicted date preference. Predictions from integration theory model. Data after Shanteau and Nagy (1976).

COMMENT BY
EDWARD JONES
PRINCETON UNIVERSITY

Edward Jones

By any standard that one wishes to apply this is an impressive document. (And the standards certainly include physical bulk). Hammond, McClelland, and Mumpower have attempted the almost impossible task of mastering the theories and paradigmatic procedures of six traditions in judgment and decision making and they have presented us with a variety of intelligent insights into their similarities and differences. As the lonely attributionist at this gathering I feel, after reading this document, like a poor relation whose cupboard contains little but two-by-two bones and scraps of ordinal measurement. In my critique I shall attempt to argue that our cupboard is not quite so bare and try to convey a sense of restiveness at having the attribution approach evaluated in terms of criteria tangential to its main purposes.

Hammond and his colleagues have been courageous in facing head-on some of the most difficult problems in measurement, theory construction, and the philosophy of science. My admiration is boundless, but my intellectual satisfaction is somewhat alloyed by the nagging suspicion that many of the comparative rubrics are rather superficial or phenotypic. The authors hope that their comparative exercise will help in the construction of a cumulative science of human judgment and decision making. I worry, perhaps without warrant, that if we are too impatient with noncomparable designs and procedures we will prematurely purchase an illusion of cumulative science at the expense of learning less about the whole range of problems we set out to study. There are times when I think a little procedural anarchy can be a good thing. The thrust of the present report may underestimate the importance of haphazard developments in different laboratories attacking slightly different problems.

Sometimes I think we have too little faith in the role of our inefficient, untidy, but ultimately triumphant scientific democracy. Agreement on standardized procedures sounds hygienic, but it may also create impossible chasms between methodological traditions.

Another problem that I find hard to resolve in trying to process the comparative statements scattered throughout the report is raised by the extent to which the approaches deal with such different ranges of concerns. Attribution theory, for example, inhabits a sprawling terrain, unashamedly sticks its nose into many other people's business, freely borrows procedures and measuring devices from related traditions, and is generally a label that conceals great heterogeneity of focus and method. The decision theory approaches, on the other hand, are definable much more in terms of a standardized paradigm and various measurement preferences. Since the emphasis in the report is on specific judgmental decisions, rather than cognitive processes in general, much of the organization seems less relevant to attribution theory than to other approaches. Also, the heterogeneity of attribution theory approaches creates problems for Hammond et al. since it is usually possible to find exceptions for every one of their generalizations about the attributional approach. I shall try to restrain myself from an over-emphasis on such exceptions, but some of them will naturally find their way into my critique.

Since I have been appointed the custodian of attribution theory, perhaps the most reasonable way to proceed is for me to enumerate the characterizations of attribution theory in the report (more or less in the order in which they appear) and to offer critical comment where I think it might be helpful.

1. II-C-10; AT is described as highly general in its aims, without noticeable boundaries, and concerned with problems ranging from interpersonal learning, through conflict, to judgments of causal locus. It concerns the psychology of common sense and its locus is the "customary habitat" of ordinary people. It "disavows little with regard to its explanatory domain." [This sounds a little like what I've just been saying, but the emphasis on an "it" makes it seem that there is some over-arching theory that attempts to explain everything. Instead, there is a handful of notions like "augmentation" and "discounting" which can be shown to operate in a variety of social explanation contexts. They by no means explain all there is to know about such contexts.]

2. There are a number of points related to the idea that AT is indiffernt to quantification. On II-E-27 this becomes elevated to a "meta-theoretical aversion." AT is destined to "provide the psychological flesh for the quantitative bones of IIT." [In general these descriptions are probably apt, for there certainly is little interest in developing a standard format for measuring perceived causality, and the independent variables of attributional experiments are still crudely manipulated for the most part. The reason for this probably lies in the priority assigned to "mundane realism" and the layman's naive phenomenology. Here we are very likely victims of the Heisenberg principle: the attribution researcher is nervous about the extent to which his questions elicit attributions that would not otherwise be there. In other words, does the form of his questioning shape the nature of the attributions described? Many attributionists would find the continual probing of the Bayesian researcher highly reactive and artificializing--perhaps carrying the subject a considerable distance away from his normal

cognitive functioning. The same can be said for many of the within-group designs of IIT. It must be acknowledged that the equation of common sense with the "application of an orthogonal components design" may be vulnerable to criticism, though of course the subject is typically exposed to only one cell of the design. Many of us feel that the avoidance of repeated measurement designs reduces the dangers of experimenter demand and consequently avoids (it is hoped) converting the subject into a psychological theorist or a spokesman for perceived consensus.]

3. AT is described along with other group II approaches as "non-prescriptive." Behavior is evaluated against empirical criteria and not against optimal choices. The report acknowledges a rare exception in an essay by Jones and McGillis, II-D-34. (Unfortunately we are here seriously misquoted: instead of saying "correspondent inference theory has a different goal than attribution theory" we in fact said, "Kelley's theory has a different goal than correspondent inference theory." Also the context is different than that implied. We were talking about entity versus person emphasis rather than the issue of prescriptive versus descriptive approaches.) [In fact there are increasing signs of AT's concern with attributional bias as compared both to rational baseline models like correspondent inference theory, and other attempts to match responses against an existing state of affairs. Thus, Ross and his colleagues emphasize the biasing effects of impression perseverance, and speak also of a "false consensus bias" in estimating the preferences of others. Nisbett and his colleagues sound very much like Tversky and Kahneman when they demonstrate our responsiveness to stereotypic cultural explanations at the expense of accuracy. Our work on attitude attribution perhaps provides the clearest example of how

a bias may be evaluated against a rational baseline model, and Kelley's ANOVA cube has been used in much the same way.]

4. AT is concerned with the search for "structurally safe"...
"stable reference points in social circumstances" (II-E-18).

[Although this is presented as evidence of a lack of emphasis on uncertainty, to say that man searches for invariance does not ignore the existence of variance as a problem.]

5. It is said (II-E-20) that AT places an implicit emphasis on processes of cognitive organization. [The record shows much more interest in the antecedents and consequences of attribution than in the processes of attribution themselves.]

6. A turn toward probabilism is noted. [Correspondent inference theory has contained a conceptual reference to probabilities for some time, but probability estimates have been used more in the establishment of expectancy strengths for independent variable manipulation than as a major dependent variable.]

7. On II-G-8 it is noted that there are "no direct efforts to apply the results of AT research to specific problems in the real world." [This strikes me as rather strong in view of the emphasis by Valins and Nisbett on correcting misattributions in diagnosis and therapy, Kelley's research on attributional conflict, and Sharon Brehm's attempt to extract clinical applications from attribution research. As an aside, it is also very apparent that the Ross and Lepper work on perseverance is directly relevant to the debriefing process in experimental social psychology.]

8. AT is nomothetic and its use of elaborate manipulations and scenarios (striving for a "concrete realistic level of presentation") makes demand characteristics a definite problem for within-subject comparisons. AT "rarely, if ever, constructs models of individual judgment or processes." [Both points are correct.]

9. We are told on page III-C-9 that the variety of tasks used by AT research is "more to provide interesting (for social psychologists) experimental contexts than to show generalizability of attribution processes over substantive areas." [One wonders why they would be "interesting to social psychologists" if they were not generalizable. This is one of the few places where a snide note creeps into the presentation.]

10. "Little attention is devoted to formal task sampling in AT", III-C-10. [Perhaps this is true, but it is not clear what is meant by "formal task". We are told (III-C-2) that formal task sampling is exemplified by "inter-correlations among cues, redundancy or inconsistency of information, type of response scale." Certainly the work on order effects would fall into this category. McArthur's work is noted as an example of formal task sampling, but this could also be viewed as substantive. On the dependent variable side, Gurwitz and Panciera have attempted to show the effects of posing different kinds of questions after a complex experimental experience.]

11. Time and again, AT is gently assailed for the complexity of its stimulus manipulations. AT research "uses the most complex, undecomposed objects" (III-D-6). It is the only one of the six approaches that "does not formally decompose judgments into weights and function forms." (III-E-5). AT proceeds instead by using "the

method of diacritical confrontation"--pitting two opposed hypotheses against each other in a judgment task. [It is probably true that there is considerably more naturalistic "embeddedness" in the stimulus materials presented to AT subjects than in the other approaches, but the complexity of the entire stimulus configuration should not be confused with the complexity of the information being manipulated. Often the independent variable involves very slight changes in but a few words of the instructional or social episodic material presented. Once again it may be noted that it is because of this that within group designs are usually impractical. Also, "diacritical confrontations" are hardly the order of the day.]

12. An "organizing principle for the environment" is not explicitly postulated in AT. Instead, AT tries to come to terms with the fact that judges impose order on the world in order to make sense out of it (IV-30). [It is certainly true that AT has moved in a different direction than SJT from their common functionalistic origins. The AT emphasis is still on functionalism, but more in the sense of the salutary importance of subjective control than in the sense of accurate attainment. It is recognized, in other words, that people will attempt to make sense out of the environment and it is this process of trying to make sense that AT explicates. That stream of AT coming out of person perception research has made investigators very wary of being trapped once again in the accuracy maze from which there is no exit (i.e., who are the most accurate person perceivers and who are the most easily judged targets). It is interesting that the original actor-observer divergence proposition tried to steer clear of deciding whether the actor or the observer was more accurate. However, subsequent writers have been less able to avoid this problematic reference to differential accuracy

(e.g., Monson and Snyder). When the report says "AT is concerned on a theoretical level with the accuracy of the individual's inferences with respect to the environmental system", IV-86, this is true only in a very long range, ultimate sense. Among other things, AT is concerned to show that the need to impose order or structure can introduce many biasing factors associated with premature encoding or categorization.]

13. On page IV-60 it is noted that AT subjects "make holistic, intuitive ratings." [One wonders, as opposed to what? Aren't probability judgments intuitive? Can they not be holistic?]

Let me close with some comments on the origins of current attribution theory research which may help to broaden the historical perspective beyond that presented in the current report. Heider still exerts a pervasive influence, if only because his writing touched on so many facets of interpersonal relations. But it helps to recognize that the current wave of attribution research primarily draws from and contributes to experimental social psychology, which has its own traditions, icons, and receptivities. The attributional approach may have begun with Heider's 1944 paper on phenomenal causality, but the roots of contemporary attribution research lie equally in Festinger's social comparison theory, Bem's self-perception theory, Schachter's theory of emotional experience, and the Tversky-Kahneman notion of laymen's heuristics. In addition, the current interest in attribution research is certainly resonant with a long-standing interest in the phenomenological approach in social psychology even though AT is not explicitly phenomenological.

Attribution theory is not aptly described as a theory. It is more like an atmosphere, a setting for certain kinds of questions,

a collection of loosely interrelated problems. Most people would include under the heading of attribution research many developments that are only tangentially attributional in the strict sense of an explicit concern with phenomenal causality. In addition to a continuing interest in (a) the questions of divergent perspectives, (b) "defensive attribution" and responses to success and failure, (c) the "fundamental attribution error", (d) dispositional inferences in person perception, and (e) the role of Kelley's ANOVA variables, there is an emerging interest in: (f) the determinants of intrinsic versus extrinsic motivation, (g) the self-fulfilling prophecy--coming to terms with expectations and labels applied by others, (h) the phenomena of objective-self awareness, (i) the perception of personality and the application of the theory of prototypes, (j) the cultural (versus ecologically valid) bases of phenomenally experienced cognitive processes, (k) positive and negative placebo effects, (l) self-presentation strategies for augmenting one's power, (m) the relative irreversibility of cognitive processes, (n) a revival of interest in stereotyping and the coding of social information, (o) a strong thread of concern with self-attribution, and last but surely not least, (p) the reattribution of emotional arousal.

APPENDIX B

Distribution List

Director, Engineering Psychology
Programs, Code 455
Office of Naval Research
800 No. Quincy Street
Arlington, VA 22217

Defense Documentation Center
Cameron Station
Alexandria, VA 22314

Dr. Stephen J. Andriole
Director, Cybernetics Tech. Office
Advanced Research Projects Agency
1400 Wilson Blvd
Arlington, VA 22209

Cdr. Paul Chatelier
OAD(E&LS) ODDR&E
Pentagon, Room 3D129
Washington, D.C. 20301

Office of Naval Research
International Programs
Code 1021P
800 No. Quincy Street
Arlington, VA 22217

Dr. Judith Daly
Cybernetics Technology Office
Defense Adv. Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Director, Naval Analysis Programs
Code 431
Office of Naval Research
800 No. Quincy Street
Arlington, VA 22217

Director, Operations Research
Programs, Code 434
Office of Naval Research
800 No. Quincy Street
Arlington, VA 22217

Director, Statistics and
Probability Program, Code 436
Office of Naval Research
800 No. Quincy Street
Arlington, VA 22217

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, D.C. 20360

Commanding Officer
ONR Branch Office
ATTN: Dr. J. Lester
495 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. Charles Davis
536 So. Clark Street
Chicago, IL 60605

Commanding Officer
ONR Branch Office
ATTN: Dr. E. Gloye
1030 E. Green Street
Pasadena, CA 91106

Dr. Bruce McDonald
Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco 96503

Director, Naval Research Laboratory
Technical Information Division
Code 2627
Washington, DC 20375

Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Department of the Navy
Washington, D.C. 20350

Dr. James Curtin
Naval Sea Systems Command
Personnel & Training Analyses Office
NAVSEA 074C1
Washington, D.C. 20362

Director
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

Distribution List

CDR R. Gibson
Bureau of Medicine & Surgery
Aerospace Psychology Branch
Code 513
Washington, D.C. 20372

Bureau of Naval Personnel
Special Assistant for Research
Liaison
PERS-OR
Washington, D.C. 20370

Navy Personnel Research and
Development Center
Management Support Department
Code 210
San Diego, CA 92152

Navy Personnel Research and
Development Center
Code 305
San Diego, CA 92152

LCDR P. M. Curran
Human Factors Engineering Branch
Crew Systems Department, Code 4021
Naval Air Development Center
Johnsville
Warminster, PA 18950

LCDR William Moroney
Human Factors Engineering Branch
Code 1226
Pacific Missile Test Center
Point Mugu, CA 93042

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-00T
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Mr. J. Barber
Headquarters, Department of the
Army, DAPE-PBR
Washington, D.C. 20546

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Edgar M. Johnson
Organization and Systems
Research Laboratory
U.S. Army Research Lab
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. John Silva
Man-System Interaction Division
Code 823, Naval Ocean Systems Center
San Diego, CA 92152

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Air University Library
Maxwell Air Force Base, AL 36112

Dr. Robert R. Mackie
Human Factors Research, Inc.
Santa Barbara Research Park
6780 Cortona Drive
Goleta, CA 93017

Dr. Gershon Weltman
Perceptronic, Inc.
6271 Varie Avenue
Woodland Hills, CA 91364

Distribution List

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Arthur Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Professor Douglas E. Hunter
Defense Intelligence School
Washington, D.C. 20374

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, D.C. 20362

Lt. Col. Joseph A. Birt
Human Engineering Division
Aerospace Medical Research Laboratory
Wright Patterson AFB, OH 45433

Dr. Clinton Kelly
Decisions & Designs, Inc.
8400 Westpark Drive, Suite 600
McLean, VA 22101

Dr. Terrence R. Mitchell
University of Washington
Seattle, WA 98195

Dr. Melvin R. Novick
University of Iowa
Lindquist Center for Measurement
Iowa City, IA 52242

Dr. Jesse Orlansky
Institute for Defense Analysis
400 Army-Navy Drive
Arlington, VA 22202

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, NW
Washington, D.C. 20036 (3 cys)

Dr. William A. McClelland
Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Professor Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Ron Howard
Stanford University
Stanford, CA 94305

Dr. Ward Edwards
Director, Social Science
Research Institute
University of Southern California
Los Angeles, CA 90007

Dr. Charles Gettys
University of Oklahoma
Department of Psychology
455 West Lindsey
Norman, OK 73069

Dr. Peter C. Goldschmidt
Policy Research Incorporated
2500 North Charles Street
Baltimore, MD 21218

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Professor Judea Pearl
University of California-Los Angeles
Engineering Systems Department
405 Hilgard Avenue
Los Angeles, CA 90024

Dr. Fred Muckler
Navy Personnel Research and
Development Center
Manned Systems Design, Code 311
San Diego, CA 92152

Director, Human Factors Wing
Defense & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsville, Toronto, Ontario
CANADA

Distribution List

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND

Dr. David Zaidel
Institute for Research in Public
Safety
University of Indiana
Bloomington, IN 47401